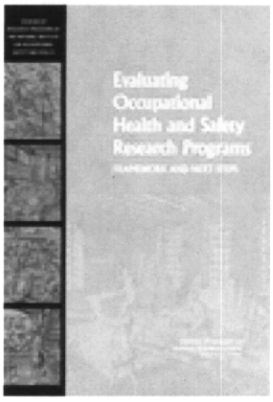

Evaluating Occupational Health and Safety Research Programs: Framework and Next Steps

Committee on the Review of NIOSH Research Programs; Institute of Medicine and National Research Council

ISBN: 0-309-13796-9, 200 pages, 6 x 9, (2009)

This free PDF was downloaded from:
http://www.nap.edu/catalog/12639.html

Visit the National Academies Press online, the authoritative source for all books from the National Academy of Sciences, the National Academy of Engineering, the Institute of Medicine, and the National Research Council:

- Download hundreds of free books in PDF
- Read thousands of books online, free
- Sign up to be notified when new books are published
- Purchase printed books
- Purchase PDFs
- Explore with our innovative research tools

Thank you for downloading this free PDF. If you have comments, questions or just want more information about the books published by the National Academies Press, you may contact our customer service department toll-free at 888-624-8373, visit us online, or send an email to comments@nap.edu.

This free book plus thousands more books are available at http://www.nap.edu.

Copyright © National Academy of Sciences. Permission is granted for this material to be shared for noncommercial, educational purposes, provided that this notice appears on the reproduced materials, the Web address of the online, full authoritative version is retained, and copies are not altered. To disseminate otherwise or to republish requires written permission from the National Academies Press.

THE NATIONAL ACADEMIES
Advisers to the Nation on Science, Engineering, and Medicine

tp://www.nap.edu/catalog/12639.html

Evaluating Occupational Health and Safety Research Programs

FRAMEWORK AND NEXT STEPS

Committee for the Review of NIOSH Research Programs

INSTITUTE OF MEDICINE *AND*
NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

THE NATIONAL ACADEMIES PRESS
Washington, D.C.
www.nap.edu

PREPUBLICATION COPY—UNCORRECTED PROOFS

tp://www.nap.edu/catalog/12639.html

THE NATIONAL ACADEMIES PRESS • 500 Fifth Street, N.W. • Washington, DC 20001

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This study was requested by the National Institute for Occupational Safety and Health of the Centers for Disease Control and Prevention and supported by Award No. 211-2006-19152 T.O. #1 between the National Academy of Sciences and the Centers for Disease Control and Prevention. Any opinions, findings, conclusions, or recommendations expressed in the publication are those of the author(s) and do not necessarily reflect the view of the organizations or agencies that provided support for this project.

International Standard Book Number-13: XXXXX
International Standard Book Number-10: XXXXX

Additional copies of this report are available from the National Academies Press, 500 Fifth Street, N.W., Lockbox 285, Washington, DC 20055; (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area); Internet, http://www.nap.edu.

For more information about the Institute of Medicine, visit the IOM home page at: **www.iom.edu.**

Copyright 2009 by the National Academy of Sciences. All rights reserved.

Printed in the United States of America

Suggested citation: IOM (Institute of Medicine) and National Research Council. 2009. *Evaluating occupational health and safety research programs: Framework and next steps.* Washington, DC: The National Academies Press.

PREPUBLICATION COPY—UNCORRECTED PROOFS

tp://www.nap.edu/catalog/12639.html

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

www.national-academies.org

PREPUBLICATION COPY—UNCORRECTED PROOFS

tp://www.nap.edu/catalog/12639.html

tp://www.nap.edu/catalog/12639.html

COMMITTEE FOR THE REVIEW OF NIOSH RESEARCH PROGRAMS*

DAVID H. WEGMAN (*Chair*), University of Massachusetts–Lowell
WILLIAM B. BUNN, III, International Truck and Engine Corporation, Warrenville, IL
CARLOS A. CAMARGO, JR., Massachusetts General Hospital, Boston
SUSAN E. COZZENS, Georgia Institute of Technology, Atlanta
LETITIA K. DAVIS, Massachusetts Department of Public Health, Boston
JAMES W. DEARING, Kaiser Permanente, Denver, CO
FRED A. METTLER, JR., University of New Mexico, Albuquerque
FRANKLIN E. MIRER, Hunter School of Health Sciences, New York
JACQUELINE NOWELL, United Food and Commercial Workers International Union, Washington, DC
RAJA V. RAMANI, Pennsylvania State University–University Park
JORMA RANTANEN, Finnish Institute of Occupational Health, Helsinki, Finland
ROSEMARY K. SOKAS, University of Illinois, Chicago (until January 2009)
RICHARD L. TUCKER, The University of Texas–Austin
JAMES J. ZUICHES, North Carolina State University–Raleigh

Study Staff

CATHY LIVERMAN, Project Director (since January 2008)
EVAN B. DOUPLE, Staff Officer (until January 2008)
SAMMANTHA MAGSINO, Program Officer
ANDREW M. POPE, Board Director
GREGORY H. SYMMES, Deputy Executive Director
JUDY ESTEP, Program Associate

---

*Maxine Hayes, University of Washington, Olympia (resigned August 2005).
Donald Henderson, State University of New York (resigned June 2005).
Joseph. S. Wholey, University of Southern California–Los Angeles (resigned March 2006).

PREPUBLICATION COPY—UNCORRECTED PROOFS

tp://www.nap.edu/catalog/12639.html

tp://www.nap.edu/catalog/12639.html

Reviewers

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process. We wish to thank the following individuals for their review of this report:

Alfred Franzblau, University of Michigan School of Public Health
Patricia A. Gleason, Safety Equipment Institute
Paul D. Gunderson, Dakota Center for Technology-Optimized Agriculture, Lake Region State College
Brian M. Kleiner, Virginia Tech
Frances P. Lawrenz, University of Minnesota
Laura C. Leviton, The Robert Wood Johnson Foundation
Julia Melkers, Georgia Institute of Technology
Jonathan G. Price, Nevada Bureau of Mines and Geology, University of Nevada
Stanley C. Suboleski, Federal Mine Safety and Health Review Commission (retired), Midlothian, Virginia

Although the reviewers listed above have provided many constructive comments and suggestions, they were not asked to endorse the

tp://www.nap.edu/catalog/12639.html

conclusions or recommendations, nor did they see the final draft of the report before its release. The review of this report was overseen by **David G. Hoel,** Medical University of South Carolina and **John C. Bailar III,** The University of Chicago, (emeritus). Appointed by the National Research Council and Institute of Medicine, they were responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and the institution.

tp://www.nap.edu/catalog/12639.html

Preface

O wad some power the giftie gie us
to see oursel's as ithers see us.[1]
— Robert Burns

In our modern society, we observe evaluation or assessment efforts in all walks of life. Among the many systematic approaches to evaluation are those seen in outcomes assessment for clinical practice, in standard exams for student performance in primary and secondary education, in accreditation reviews for higher education programs and institutions as well as hospitals, and in the Government Performance and Results Act and Performance Assessment Rating Tool approaches to assess effectiveness of the great variety of federal programs. Each of these settings has promoters and detractors of the objectives to be achieved by assessment as well as of the specific assessment mechanisms used for the evaluation.

Hardly anyone anticipates the prospect of being evaluated with pleasure. Most, however, agree that the often-burdensome process of preparing for, as well as undergoing, an evaluation ends up as instructive and promotes improved individual or program performance. Facing the challenges of any assessment by an external organization entails assem-

---

[1]Robert Burns, *To A Louse*. On seeing one on a lady's bonnet at church.
Original:

O wad some power the giftie gie us
To see oursel's as ithers see us!
It wad frae monie a blunder free us,
An' foolish notion:
What airs in dress an' gait wad lea'e us,
An' ev'n devotion!

Standard English translation:

O would some Power the gift to give us
To see ourselves as others see us!
It would from many a blunder free us,
And foolish notion:
What airs in dress and gait would leave us,
And even devotion!

PREPUBLICATION COPY—UNCORRECTED PROOFS

tp://www.nap.edu/catalog/12639.html

bling and organizing large quantities of data and program details, frequently accompanied by the task of preparing a self-assessment.

When the National Institute for Occupational Health and Safety (NIOSH) determined that a formal external program evaluation would contribute importantly to achieving its mission, the Institute took the bold step of asking that the evaluation be not only qualitative, but also quantitative. Furthermore, despite the predominant research mission of the Institute, the leadership determined that each program's activities and outputs should be assessed on a standard that judged (1) whether they were contributing to new scientific knowledge, and (2) to what extent the program was making a significant and important contribution to improving worker safety and health. The Institute, thus, wished to acknowledge fully the applied nature of its research mission.

As the framework committee began this task in 2004, there were many unknowns about what the challenges would be in providing consistent evaluations across a set of widely diverse NIOSH programs. The evaluation framework developed by the committee has now been used by eight separate evaluation committees and has been found to be a useful and thorough approach for program evaluation. The resulting evaluations have provided a consistent assessment and have identified a number of recommendations to improve the NIOSH programs. In general the separate evaluations found the programs to be productive, to represent substantial quality and applicability, and to be efficient in using very limited resources.

When the eight program evaluations were examined together, some common needs emerged that would benefit future evaluations and NIOSH as a whole. Most of the evaluation committees noted the lack of adequate surveillance data on occupational exposures, illnesses, and injuries and made recommendations to bolster surveillance systems. An increased focus on strategic planning was another recurring theme, as was the need to strengthen efforts to move research findings to the worksite. Improvements in integrating extramural and intramural research at NIOSH were also recommended.

This report provides an opportunity to step back and look at the broader picture of the evaluation process developed for the NIOSH research programs. The framework committee has heard throughout the process about the strengths and challenges of this endeavor. In November 2008 the framework committee held a workshop in which that committee's members, members of the eight evaluation committees, NIOSH staff, and National Academies' staff met to discuss the lessons learned.

tp://www.nap.edu/catalog/12639.html

Thus, this report draws on those insights as well as others gained by the framework committee over the course of the past 4 years.

The committee and the individual evaluation committees greatly benefited from the thorough briefings and informative discussions with NIOSH staff members. On behalf of the committee, I want to especially thank Ray Sinclair and Lewis Wade, who provided excellent guidance from the inception of the project and who held fast to the goals of examining the relevance and impact of NIOSH's work on the end outcomes of worker safety and health, an evaluation process envisioned by John Howard. This committee and each of the evaluation committees appreciate the dedicated efforts by NIOSH staff in compiling information and responding to numerous inquiries. It is heartening to see that NIOSH staff have found that the evaluations address their needs and are devoting similar efforts to the development of implementation plans in response to the evaluations.

Chairing this National Academies committee and having the opportunity to interact with the eight evaluation committees was a privilege and a pleasure. The framework committee members sustained their energy, interest, and dedication to this task over the course of four years. In addition to serving on this committee, many members also served as liaisons or members of the evaluation committees, and I thank them for the time and high level of engagement they gave to this evaluation process.

NIOSH has a large task—conducting research to improve occupational safety and health—but limited resources. This evaluation process has shown the great relevance and impact of NIOSH's work. Our hope is that all of these efforts will contribute to further improvements in the health and safety of workers.

David Wegman, *Chair*
Committee for the Review of
NIOSH Research Programs

tp://www.nap.edu/catalog/12639.html

tp://www.nap.edu/catalog/12639.html

Contents

SUMMARY 1

1 INTRODUCTION 11
Scope of the Task, 12
Overview of NIOSH, 14
Developing and Implementing the Evaluation Framework, 17
Overview of this Report, 24
References, 25

2 THE PROGRAM EVALUATION CONTEXT 27
Program Evaluation, 27
Logic Models, 29
Role of Stakeholders, 31
Methods of Evaluation, 32
Summary, 37
References, 37

3 EVALUATION FRAMEWORK 41
Overview of the Evaluation Framework, 42
Evaluation Committees, 47
Steps in the Evaluation Process, 48
References, 81

4 IMPROVING THE EVALUATION PROCESS 83
Evaluation Framework, 84
Composition of the Evaluation Committees, 89
Agency Inputs to the Evaluation, 89

tp://www.nap.edu/catalog/12639.html

Stakeholder Input, 96
Time Lines for Evaluations, 97
Evaluation Committee Reports and Recommendations, 98
Summary, 100
References, 101

5 RECOMMENDATIONS FOR MOVING FORWARD 103
Ongoing Evaluation, 103
Continue to Bolster Research Translation Efforts, 105
Enhance Occupational Health and Safety Surveillance, 107
Integrate Evaluations of Extramural and Intramural Research, 109
On the Horizon, 110
References, 112

APPENDIXES
A Meeting Agendas 115
B Biographical Sketches of Committee Members 123

tp://www.nap.edu/catalog/12639.html

Abbreviations and Acronyms

AFF	Agriculture, Forestry, and Fishing
AOEC	Association of Occupational and Environmental Clinics
BLS	Bureau of Labor Statistics
BSC	Board of Scientific Counselors
CDC	Centers for Disease Control and Prevention
CSTE	Council of State and Territorial Epidemiologists
DoD	U.S. Department of Defense
EPA	Environmental Protection Agency
FACE	Fatality Assessment and Control Evaluation Program
GPRA	Government Performance and Results Act
HHE	Health Hazard Evaluation
HHS	U.S. Department of Health and Human Services
IOM	Institute of Medicine
MSHA	Mine Safety and Health Administration

PREPUBLICATION COPY—UNCORRECTED PROOFS

tp://www.nap.edu/catalog/12639.html

NIH	National Institutes of Health
NIOSH	National Institute for Occupational Safety and Health
NORA	National Occupational Research Agenda
NRC	National Research Council
OMB	Office of Management and Budget
OSHA	Occupational Safety and Health Administration
PART	Performance Assessment Rating Tool
PEL	permissible exposure limit
PPT	Personal Protective Technology
RFA	Requests for Applications
USDA	U.S. Department of Agriculture

tp://www.nap.edu/catalog/12639.html

Summary

ABSTRACT *The goal of research conducted by the National Institute for Occupational Safety and Health (NIOSH) is protecting and improving the health and safety of workers. The Institute of Medicine and the National Research Council conducted a series of evaluations of NIOSH research programs that assessed the relevance and impact of NIOSH's work on meeting this goal. Eight NIOSH programs were evaluated based on a common framework. This report provides the lessons learned in the evaluation process, a revised evaluation framework, and the following recommendations for evaluating occupational health and safety research programs: continue systematic external evaluations, bolster research translation efforts, enhance occupational health and safety surveillance, and integrate evaluations of intramural and extramural research.*

Preventing tractor rollovers on farmers, protecting construction workers from falls, improving the health of miners in dusty environments, reducing back injuries in nursing aides, developing substitute materials to eliminate hazardous chemical exposures, and designing work conditions to reduce fatigue and stress are among the issues critical to improving worker safety and health. In 2007, 5,657 fatal work injuries occurred in the United States, along with an estimated 4 million nonfatal occupational injuries and illnesses among workers in private industry. In addition to the human suffering involved, these statistics are also associated with high economic costs. One estimate puts the costs of occupational injury and illnesses for all industries for 2005 at more than $160 billion. The continued attention to further improve occupational health and safety through research is not only fully warranted but such research

PREPUBLICATION COPY—UNCORRECTED PROOFS

tp://www.nap.edu/catalog/12639.html

requires critical evaluation for its relevance and impact. The core mission of the National Institute for Occupational Safety and Health (NIOSH) is to conduct research to improve and protect the health and safety of workers.

In September 2004, NIOSH contracted with The National Academies to conduct a series of evaluations of individual NIOSH research programs. This set of independent evaluations focused on the relevance and impact of each of eight NIOSH programs on reducing work-related injuries, illnesses, and hazardous exposures. From the outset of the evaluations, NIOSH leadership established their primary goal as program improvement but the context for the evaluations also included the PART (Program Assessment Rating Tool) federal agency evaluation process.

The first step in this multiphase effort was the appointment of a committee to develop an evaluation framework that was then used by eight separately appointed evaluation committees to assess NIOSH programs in hearing loss; mining; agriculture, forestry, and fishing; respiratory diseases; personal protective technology; traumatic injury; construction; and health hazard evaluation. Individual reports were produced by each evaluation committee.

At the conclusion of the eight studies, the framework committee held a public workshop in November 2008, "Evaluating NIOSH Programs: Lessons Learned and Next Steps," where the discussions focused on the experiences gained in the evaluation process with NIOSH program and senior staff, members of the NIOSH Board on Scientific Counselors, evaluation committee members, and National Academies' staff. This report provides the evaluation framework developed, implemented, and refined over the course of four years and eight program evaluations. The evaluation framework may prove applicable in evaluating other federal agency research programs. This report has two goals: (1) to summarize the evaluation process and lessons learned in the development and use of the framework and (2) to provide recommendations for future evaluation efforts.

EVALUATION FRAMEWORK

After examining different approaches to program evaluation, the framework committee decided to define the scope and stages of the evaluation process based on the logic model, a model that is widely used in program evaluation and planning.

tp://www.nap.edu/catalog/12639.html

The logic model organizes the program and its efforts into inputs (e.g., budget, staffing, facilities), activities (e.g., research studies, surveillance, exposure measurement), outputs (e.g., reports, publications, conferences, training, patents), and outcomes (e.g., collaborations, policy changes, reductions in exposures and injuries). The evaluation framework (outlined in Box S-1) developed for the evaluation of the NIOSH research programs provides criteria for assessing each component of the logic model.

As requested by NIOSH, the charge to the National Academies included scoring each program (using integer rating scales of 1 to 5) on the relevance and on the impact of the NIOSH program in improving worker safety and health. To provide guidance on evaluating and scoring these measures after assessing each NIOSH program, the framework committee developed criteria and specific questions to be used in the assessment of each component of the logic model. Assessment of strategic goals and objectives, inputs, activities, and outputs largely defined the relevance of the program; the committee examined the adequacy of the inputs and the scope and targeting of the activities and outputs in achieving the program's goals. Assessment of the intermediate and end outcomes largely defined the program's impact. The evaluation framework also included specific sets of scoring criteria for rating the program's relevance and impact on reducing work-related injuries, illnesses, or hazardous exposures (Boxes S-2 and S-3).

Setting the metric for program success at demonstrating an impact on end outcomes is laudable. Evaluation and framework committee members give NIOSH a great deal of credit for holding their research programs accountable for real outcomes that affect life and health. Evaluations of research supported by other federal agencies often focus on output productivity and intermediate outcomes and do not hold the agency accountable for real-world impacts. This may be appropriate for the goals of some research programs, but NIOSH's focus is on applied research and thus the end outcomes are appropriately focused on improving worker safety and health. The evaluation framework acknowledges that actions of those in industry, labor, regulatory entities, and others beyond NIOSH's control are also necessary for NIOSH programs to have impacts.

The challenge for the evaluation committees was to search for data on end outcomes while also thoroughly examining the multiple pathways leading to intermediate outcomes and from there to end outcomes. NIOSH staff and the evaluation committees used the logic model to

tp://www.nap.edu/catalog/12639.html

document a wide range of intermediate outcomes. Particularly in the absence of good end outcome data, a frequent finding, attention was often focused on the most observable pathways between program activities and end outcomes.

The scoring criteria for relevance and impact are provided in Boxes 2 and 3. The criteria for assessment of relevance are focused on determining whether the program appropriately sets priorities among research needs as well as how engaged the program is in appropriate transfer activities to move research findings into the workplace. The criteria for assessment of impact are focused on completed work and the extent to which the work has directly or indirectly led to improvement in worker health and safety. The scoring systems serve as a starting point for descriptive text that explains what the score meant and describes the program's strengths and limitations that led to the score. Although there was initial concern that the scores would be the only endpoint noted from the evaluation reports, the framework committee is pleased to see the detailed strategic plans and action plans that are being developed and implemented in NIOSH's response to the recommendations.

BOX S-1
Steps in the Evaluation Process

1. Gaither appropriate information
2. Assess external factors.
3. Identify time frame to be evaluated.
4. Identify major occupational safety and health challenges in program area.
5. Analyze program goals and objectives.
6. Identify major program components.
7. Evaluate program inputs, activities, outputs, and outcomes.
8. Determine scores for relevance and impact and provide the rationale.
9. Assess the NIOSH process for targeting priority research needs and provide the committee's assessment of emerging issues.
10. Prepare report by using the template provided as a guide.

tp://www.nap.edu/catalog/12639.html

BOX S-2
Scoring Criteria for Relevance

5 = Research is in high-priority subject areas and the NIOSH program is significantly engaged in appropriate transfer activities for completed research projects/reported research results.

4 = Research is in high-priority subject areas and the NIOSH program is engaged in appropriate transfer activities for completed research projects/reported research results; or research is in priority subject areas and the NIOSH program is significantly engaged in appropriate transfer activities for completed research projects/reported research results.

3 = Research is in high-priority subject areas, but the NIOSH program is not engaged in appropriate transfer activities; or research is in priority subject areas but the NIOSH program is not significantly engaged in appropriate transfer activities; or research focuses on lesser priorities but the NIOSH program is significantly engaged in appropriate transfer activities.

2 = Research program is focused on lesser priorities and the NIOSH program is not significantly engaged in appropriate transfer activities.

1 = Research program is not focused on priorities.

BOX S-3
Scoring Criteria for Impact

5 = Research program has made major contribution(s) to worker health and safety on the basis of end outcomes or well-accepted intermediate outcomes.

4 = Research program has made some contributions to end outcomes or well-accepted intermediate outcomes.

3 = Research program activities are ongoing and outputs are produced that are likely to result in improvements in worker health and safety. Well-accepted outcomes have not been recorded.

2 = Research program activities are ongoing and outputs are produced that may result in new knowledge or technology, but only limited application is expected. Well-accepted outcomes have not been recorded.

1 = Research activities and outputs do not result in or are not likely to have any application.

NA = Impact cannot be assessed; program is not mature enough.

tp://www.nap.edu/catalog/12639.html

DATA FOR THE EVALUATION PROCESS

NIOSH provided detailed evidence packages to each evaluation committee with information on the program including program descriptions, staffing levels, program goals and objectives, and details on the program and its accomplishments. Although the evidence package was the core input to each evaluation, it was one of many sources of information that the evaluation committees assessed. Other sources included committee requests to NIOSH for additional information, presentations by NIOSH staff and academic researchers, stakeholder presentations and other input, and in some cases, site visits. Input from organizations and individuals with an interest in the mission of the program was vital to the evaluation process, particularly given the limited surveillance and other end-outcome data related to determining the program's impact on reducing hazardous occupational exposures and worker injuries and illnesses.

IMPROVING THE EVALUATION PROCESS

Throughout the course of the eight evaluations, the framework and evaluation committees exchanged information on the strengths and limitations of the framework and the evaluation process. A workshop at the conclusion of the eight studies provided an opportunity for NIOSH staff, National Academies' staff, and evaluation and framework committee members to reflect on ways to improve the process. The report suggests a number of areas in which the evaluation process could be improved, including the provision of more information in the evidence packages on priority setting efforts and on budget and staffing; additional input on extramural research and the connections between the intramural and extramural program objectives; ensuring plenty of opportunities for input by external stakeholders and agency staff; considerations regarding the time line for the evaluations; and options of other types of recommendations or ways of categorizing recommendations in the evaluation reports. Increased opportunities, particularly informal opportunities, for NIOSH staff to discuss issues with evaluation committees would be helpful as would increased attention by evaluation committees to some of the more indirect measures of intermediate outcomes. The committee also recognized that logic models are quite linear and focus on readily observable short and medium term outcomes. Evaluation committees, therefore, need to be open to exploring less linear aspects of knowledge develop-

tp://www.nap.edu/catalog/12639.html

ment and flow. Concerted efforts need to be made to include the more diffuse contributions of the program to the development of general knowledge and human capital in the field of occupational safety and health.

RECOMMENDATIONS

Drawing on the lessons learned in developing the evaluation framework and in applying the framework in eight program evaluations, the following recommendations provide the framework committee's thoughts on moving forward in program evaluation, particularly from a long-range perspective. These recommendations may also be informative for other federal agency program evaluations.

Ongoing Evaluation

Evaluation of research programs at regular intervals has become the norm, with the trend toward internal management reviews supplemented by periodic evaluation by external parties. Competently done external evaluation removes the unconscious bias of managers with regard to their programs; takes organizational competition out of the assessment; and usually provides new insights while reinforcing some of what managers already knew, but could not act on. Whatever the composition or structure of external review, the research program should conduct a self-study prior to external review. This is analogous to the self-study an academic program would conduct prior to accreditation.

Recommendation 1 ***Continue Systematic External Evaluations***
NIOSH should establish a system for periodic external evaluation complemented by internal self-assessments on a regular basis. Program or agency-wide evaluations should begin with strong self-evaluation efforts that allow the program or agency to assemble and analyze data and act on relevant findings concerning the program's strengths, weaknesses, and opportunities.

tp://www.nap.edu/catalog/12639.html

Continue to Bolster Research Translation Efforts

From the perspective of NIOSH, research transfer or translation can be seen to encompass activities that staff and researchers engage in to increase the likelihood that results of research will be used to improve worker safety and health as well as studies of research translation conducted or funded to increase knowledge about which approaches are most effective. Research transfer is a commendable new emphasis in the agency, and one that the evaluation committees noted has provided a number of positive intermediate outcomes. The committee believes NIOSH has a role to play not just in demonstrating and testing research-to-practice approaches, but also in documenting and testing its inverse, practice-to-research. Often the most effective research translation occurs through iterative learning. Practitioners can learn from researchers, but it is at least as important for researchers to learn from practitioners so that the new knowledge, practices, programs, and technologies that researchers create are informed by real-world workplace conditions.

Recommendation 2 ***Continue to Build and Improve Research Translation Efforts***

NIOSH should continue to build and improve its research translation efforts with an emphasis on:

- **ongoing assessment and improvement of its research translation efforts through formative evaluation processes of listening to those in the workplace (workers and employers) and beyond (product designers, architects, health care providers, etc.), both to identify intervention needs and to provide early feedback regarding research translation products to improve the interventions; and**
- **building the capacity to implement and evaluate research translation efforts, both as research-to-practice and as practice-to-research.**

tp://www.nap.edu/catalog/12639.html

Enhance Occupational Health and Safety Surveillance

The logic model approach to evaluation—used as the basis for the framework presented in this report—relies heavily on surveillance data on health outcomes and workplace exposures to evaluate strategic priorities and assess program impact. Surveillance data are also critical program inputs, and the extent to which research programs have considered surveillance data in setting research priorities is an important determinant of program relevance. Although a comprehensive system for tracking fatal occupational injuries in the United States is in place, the current approaches to surveillance of occupational illnesses and nonfatal occupational injuries are fragmented and incomplete, and only limited surveillance data on exposure to hazards are available. Surveillance is a necessity for monitoring long-term progress in reducing hazardous exposures and work-related injuries and illnesses.

Recommendation 3 ***Increase and Improve Surveillance to Benchmark Progress***

NIOSH should increase and improve surveillance of work-related injuries, illnesses, exposures, and working conditions so that information needed to assess program relevance and impact will be available for future evaluations. Enhanced surveillance should prove informative in balancing research priorities.

Integrate Evaluations of Extramural and Intramural Research

Obtaining the full picture of NIOSH's work in a specific area of research requires examining the relevant intramural and the extramural research. However, the evaluation committees found that the extent to which the intramural and extramural components at NIOSH are currently separated makes it difficult to conduct such an assessment. Several of the evaluation committees noted a disconnect between the intramural and extramural programs. Although the committee fully supports external scientific review to determine merit for funding investigator-initiated research, the evaluation committees noted that few avenues are currently available by which NIOSH staff can provide intramural input into the development of priorities for extramural research.

tp://www.nap.edu/catalog/12639.html

Recommendation 4 *Integrate Evaluations of Intramural and Extramural Research*

Future evaluations should systematically consider intramural and extramural research activities, in terms of both evaluating the impact and relevance of each type of research and assessing the extent to which intramural and extramural research are integrated in strategic planning.

BOX S- 4
Recommendations

Recommendation 1 ***Continue Systematic External Evaluations***
NIOSH should establish a system for periodic external evaluation complemented by internal self-assessments on a regular basis. Program or agency-wide evaluations should begin with strong self-evaluation efforts that allow the program or agency to assemble and analyze data and act on relevant findings concerning the program's strengths, weaknesses, and opportunities.

Recommendation 2 ***Continue to Build and Improve Research Translation Efforts***
NIOSH should continue to build and improve its research translation efforts with an emphasis on:

- ongoing assessment and improvement of its research translation efforts through formative evaluation processes of listening to those in the workplace (workers and employers) and beyond (product designers, architects, health care providers, etc.), both to identify intervention needs and to provide early feedback regarding research translation products to improve the interventions; and
- building the capacity to implement and evaluate research translation efforts, both as research-to-practice and as practice-to-research.

Recommendation 3 ***Increase and Improve Surveillance to Benchmark Progress***
NIOSH should increase and improve surveillance of work-related injuries, illnesses, exposures, and working conditions so that information needed to assess program relevance and impact will be available for future evaluations. Enhanced surveillance should prove informative in balancing research priorities.

Recommendation 4 ***Integrate Evaluations of Intramural and Extramural Research***
Future evaluations should systematically consider both intramural and extramural research activities, in terms of both evaluating the impact and relevance of each type of research and assessing the extent to which intramural and extramural research are integrated in strategic planning.

tp://www.nap.edu/catalog/12639.html

1

Introduction

Preventing tractor rollovers on farmers, protecting construction workers from falls, improving the health of miners in dusty environments, reducing back injuries in nursing aides, developing substitute materials to eliminate hazardous chemical exposures, and designing work conditions to reduce fatigue and stress are among the issues critical to improving worker safety and health. In 2007, 5,657 fatal work injuries occurred in the United States, along with an estimated 4 million nonfatal occupational injuries and illnesses among workers in private industry (BLS, 2008, 2009). In addition to the human suffering involved, these statistics are associated with high economic costs. One estimate puts the costs of occupational injury and illnesses for all industries for 2005 at more than $160 billion (Leigh, 2008). The continued attention to further improve occupational health and safety through research is not only fully warranted, but such research requires critical evaluation of its relevance and impact. The core mission of the National Institute for Occupational Safety and Health (NIOSH) is to conduct research to improve and protect the health and safety of workers.

In September 2004, NIOSH contracted with The National Academies to conduct a series of evaluations of individual NIOSH research programs.[1] This set of independent evaluations focused on the relevance and impact of each of eight NIOSH programs on reducing work-related injuries, illnesses, and hazardous exposures. From the outset of the evaluations, NIOSH leadership established the primary goal as program

[1]A program is defined as a set of inputs and activities directed toward one or more common goals, typically under the direction of a manager or management team.

PREPUBLICATION COPY — UNCORRECTED PROOFS

tp://www.nap.edu/catalog/12639.html

improvement, but the context for the evaluations also included the PART (Program Assessment Rating Tool) federal agency evaluation process.

The first step in this multiphase effort was to develop an evaluation framework that could be applied across the set of program evaluations to enhance cross-study consistency. An Institute of Medicine (IOM)/National Research Council (NRC) committee (the framework committee) was appointed to develop the evaluation framework. The resulting evaluation framework was then used by eight separately appointed ad hoc committees (evaluation committees) to assess NIOSH programs in hearing loss; mining; agriculture, forestry, and fishing; respiratory diseases; personal protective technology; traumatic injury; construction; and health hazard evaluation. Each evaluation committee produced an individual report (IOM and NRC, 2006, 2008, 2009; NRC and IOM, 2007, 2008a,b, 2009a,b).

This report provides the evaluation framework developed, implemented, and refined over the course of four years and eight evaluations. The framework uses a standard tool in program management and evaluation—the logic model—and provides details on the types of information that are needed and questions to be considered in each phase of the evaluation. This report has two goals: (1) to summarize the evaluation process and lessons learned in the development and use of the framework and (2) to provide recommendations for future evaluation efforts. The evaluation framework may prove applicable in evaluating other federal agency research programs.

SCOPE OF THE TASK

The framework and evaluation committees followed the same basic statement of task (Box 1-1). Although the statement of task was modified to clarify specific issues or to accommodate programs that were not specifically research programs,[2] the basic objectives for the program evaluations remained the same:

[2]The charges to the committees to evaluate the NIOSH Health Hazard Evaluation and Personal Protective Technology programs were each slightly modified to accommodate the unique standards-setting and investigative aspects of these programs.

tp://www.nap.edu/catalog/12639.html

- An assessment of the relevance and impact of the NIOSH program's contribution to reducing work-related hazardous exposures, illnesses, and injuries based on integer scales of 1 to 5, with text to support the rating;
- Assessment of the program's effectiveness in targeting new research areas and identification of emerging issues that the program should be prepared to address; and
- Recommendations for program improvement.

BOX 1-1
Review of NIOSH Research Programs

Statement of Task

In response to a request from the National Institute for Occupational Safety and Health (NIOSH), the Institute of Medicine and the Division of Earth and Life Studies of The National Academies are conducting a series of evaluations of NIOSH research programs. Each evaluation is being conducted by an ad hoc committee, using a methodology and framework developed by the Committee to Review NIOSH Research Programs (framework committee).

Each evaluation committee will review the program's impact, relevance, and future directions. The evaluation committee will evaluate not only what the NIOSH research program is producing, but will also determine whether it is appropriate to credit NIOSH research with changes in workplace practices, hazardous exposures, and/or occupational illnesses and injuries, or whether the changes are the result of other factors unrelated to NIOSH.

The program reviews should focus on evaluating the program's impact and relevance to health and safety issues in the workplace and make recommendations for improvement. In conducting the review, the evaluation committee will address the following elements:

1. Assessment of the program's contribution through occupational safety and health research to reductions in workplace hazardous exposures, illnesses, or injuries through:
   a. An assessment of the relevance of the program's activities to the improvement of occupational safety and health; and
   b. An evaluation of the impact that the program's research has had in reducing work-related hazardous exposures, illnesses, and injuries.

The evaluation committee will rate the performance of the program for its relevance and impact using an integer score of 1 to 5. Impact may be assessed directly (e.g., reductions in illnesses or injuries) or, as necessary,

tp://www.nap.edu/catalog/12639.html

using intermediate outcomes to estimate impact. Qualitative narrative evaluations will be included to explain the numerical ratings.

2. Assessment of the program's effectiveness in targeting new research areas and identifying emerging issues in occupational safety and health most relevant to future improvements in workplace protection. The committee will provide a qualitative narrative assessment of the program's efforts and suggestions about emerging issues that the program should be prepared to address.

In 2008, NIOSH requested that the Committee on the Review of NIOSH Research Programs prepare a brief report to document the lessons learned in this evaluation effort and set forth the process that it recommends be used for future evaluations of occupational health and safety research programs. The report will build on the latest version of the framework used to guide the work of the eight evaluation committees completed to date, and may revise the framework as necessary based on lessons learned about the evaluation process during these reviews. The committee will draw on the evaluation literature to provide additional context for the report. Recommendations for future program evaluations or more generalizable guiding principles for these types of efforts may be provided. Input for this report will involve discussions with NIOSH staff and with chairs and committee members of the evaluation committees.

OVERVIEW OF NIOSH

Created in 1970 by the Occupational and Safety Health Act, NIOSH is charged with the responsibility to "conduct . . . research, experiments, and demonstrations relating to occupational safety and health" and to develop "innovative methods, techniques, and approaches for dealing with occupational safety and health problems" (Public Law 91–596). The focus of NIOSH's responsibilities is occupational health and safety research, along with professional education and training. NIOSH is also involved in surveillance and in providing advice to the Secretary of Labor regarding standards needed to protect workers as well as other efforts that support research and the transfer of research into the workplace. NIOSH does not have the authority to establish or enforce regulations on workplace safety and health.[3] Regulatory and enforcement authority at the federal level rests largely with the Occupational Safety and Health

[3]NIOSH has responsibility for certifying most types of occupational respirators and for updating the federal respirator certification and testing regulations (42 C.F.R. Part 84).

tp://www.nap.edu/catalog/12639.html

Administration and the Mine Safety and Health Administration, both in the Department of Labor.

NIOSH was established as an agency within the Department of Health and Human Services (HHS). Originally the director of the institute reported to the Secretary of HHS. In the 1980s, NIOSH was reorganized as one of seven components of the Centers for Disease Control and Prevention. NIOSH offices are located across the country: Washington, DC; Pittsburgh, PA; Spokane, WA; Cincinnati, OH; Atlanta, GA; and Morgantown, WV. Because NIOSH deals with issues specific to particular work sectors (e.g., mining, construction, agriculture) as well as cross-cutting issues that span multiple sectors (e.g., hearing loss, personal protective technologies, respiratory diseases), many programs now use a matrix-management approach. (For discussion of the pros and cons of matrix management, see IOM and NRC, 2006.)

The NIOSH mission is "to provide national and world leadership to prevent work-related illness, injury, disability, and death by gathering information, conducting scientific research, and translating the knowledge gained into products and services." To fulfill its mission, NIOSH has established three strategic goals (NIOSH, 2009a):

- Goal 1: Conduct research to reduce work-related illnesses and injuries.
  - Track work-related hazards, exposures, illnesses, and injuries for prevention.
  - Generate new knowledge through intramural and extramural research programs.
  - Develop innovative solutions for difficult-to-solve problems in high-risk industrial sectors.
- Goal 2: Promote safe and healthy workplaces through interventions, recommendations, and capacity building.
  - Enhance the relevance and utility of recommendations and guidance.
  - Transfer research findings, technologies, and information into practice.
  - Build capacity to address traditional and emerging hazards.
- Goal 3: Enhance global workplace safety and health through international collaborations.
  - Take a leadership role in developing a global network of occupational health centers.

tp://www.nap.edu/catalog/12639.html

- Investigate alternative approaches to workplace illness and injury reduction and provide technical assistance to implement solutions.
- Build global professional capacity to address workplace hazards through training, information sharing, and research experience.

In 1994, NIOSH embarked on a national partnership effort to identify research priorities to guide occupational health and safety research for the next decade (NIOSH, 2009b). Participants included external stakeholders[4] from many areas, including universities, large and small businesses, professional societies, government agencies, and worker organizations. The result of these efforts was the National Occupational Research Agenda (NORA). Through the NORA effort, 21 priority research areas were identified and research agendas developed. NORA was intended not only for use by NIOSH, but also for the entire occupational health and safety community. In the second decade of NORA, NIOSH continues to work with its partners to update the research agenda using an approach based on industry sectors, with an increased emphasis on moving research results into practice in the workplace. The current NIOSH portfolio is organized into 8 sector programs and 24 crosscutting programs (Box 1-2).

BOX 1-2
NIOSH Program Portfolio

The National Institute for Occupational Safety and Health (NIOSH) Program Portfolio has been organized into 8 National Occupational Research Agenda Sector Programs that represent industrial sectors, and 24 cross-sector programs.

Sector Programs:

- Agriculture, Forestry, and Fishing
- Construction
- Health Care and Social Assistance
- Manufacturing
- Mining; Oil and Gas Extraction Subsector

[4]Stakeholders are defined as the broad group of individuals or organizations with an interest in the mission of the program. Partners are the subset of stakeholders that contribute directly to program activities.

tp://www.nap.edu/catalog/12639.html

- Services
- Transportation, Warehousing, and Utilities
- Wholesale and Retail Trade

Cross-Sector Programs:

- Authoritative Recommendations
- Cancer, Reproductive and Cardiovascular Diseases
- Communications and Information Dissemination
- Economics
- Emergency Preparedness and Response
- Engineering Controls
- Exposure Assessment
- Global Collaborations
- Health Hazard Evaluation
- Hearing Loss Prevention
- Immune and Dermal Diseases
- Musculoskeletal Disorders
- Nanotechnology
- Occupational Health Disparities
- Personal Protective Technology
- Prevention Through Design
- Radiation Dose Reconstruction
- Respiratory Diseases
- Small Business Assistance and Outreach
- Surveillance
- Training Grants
- Traumatic Injury
- Work Organization and Stress-Related Disorders
- Worklife Initiative

SOURCE: NIOSH (2009c).

DEVELOPING AND IMPLEMENTING THE EVALUATION FRAMEWORK

As noted above, before embarking on a series of program evaluations, The National Academies formed a committee to develop an evaluation framework and to continue over the course of the evaluations to monitor the use of the framework and adapt it as needed to make the evaluations more effective.

tp://www.nap.edu/catalog/12639.html

Initial Framework Document

The framework committee appointed by The National Academies consisted of 14 members with expertise in occupational medicine and health, industrial health and safety, industrial hygiene, epidemiology, civil and mining engineering, sociology, program evaluation, communication, toxicology, international occupational health, and industry and workforce issues.

During the initial development of the evaluation criteria and process, the framework committee drew from the program evaluation literature and discussed evaluation strategies and precedents with NIOSH leadership and staff, industry leaders, RAND Corporation staff,[5] occupational safety organizations, labor representatives, and staff from other federal agencies (Appendix A).

The framework committee decided to use the logic model as the basis for the evaluation framework. The logic model organizes the program and its efforts into inputs (e.g., budget, staffing, facilities), activities (e.g., research studies, surveillance, exposure measurement), outputs (e.g., reports, conferences, training, patents), and outcomes (e.g., collaborations, policy changes, reductions in injuries and hazardous exposures) (see Box 1-3 for definitions and additional examples). As will be further discussed in Chapter 2, logic models are widely used in program evaluation and planning.

To provide guidance on evaluating and scoring the relevance and impact of each NIOSH program, the framework committee developed criteria and specific questions to be used in the assessment of each component of the logic model. Assessment of strategic goals and objectives, inputs, activities, and outputs largely defined the relevance of the program; the committee examined the adequacy of the inputs and the scope and targeting of the activities and outputs in achieving the program's goals. Assessment of the intermediate and end outcomes largely defined the program's impact. Overlap necessarily occurred between the assessment of relevance and impact, particularly in the assessment of information transfer.

The framework committee met three times to develop the first version of the evaluation framework, which was released in December 2005. This version of the framework was then used by the evaluation committees exam-

[5]NIOSH contracted with RAND Corporation for the development of the logic models specific to each program and assistance on identifying and assembling the materials needed for the evaluations.

tp://www.nap.edu/catalog/12639.html

ining the first four NIOSH programs to be considered: Hearing Loss Prevention; Mining; Agriculture, Forestry, and Fishing; and Respiratory Diseases.

BOX 1-3
Logic Model Terms and Examples

Planning Inputs: Stakeholder input, surveillance and intervention data, and risk assessments (e.g., input from Federal Advisory Committee Act panels or the National Occupational Research Agenda research partners, intramural surveillance information, Health Hazard Evaluations [HHEs]).

Production Inputs: Intramural and extramural funding, staffing, management structure, and physical facilities.

Activities: Efforts and work of the program, staff, grantees, and contractors (e.g., surveillance, health effects research, intervention research, health services research, information dissemination, training, technical assistance).

Outputs: Direct products of NIOSH programs that are logically related to the achievement of desirable and intended outcomes (e.g., publications in peer-reviewed journals, recommendations, reports, website content, workshops and presentations, databases, educational materials, new technologies, patents, technical assistance).

Intermediate Outcomes: Actions by stakeholders in response to NIOSH products or efforts (e.g., policy changes; production of standards or regulations; adoption of NIOSH-developed technologies; use of publications, technologies, methods, or recommendations by workers, industry, and occupational safety and health professionals in the field; citations of NIOSH research by industry and academic scientists).

End Outcomes: Improvements in safety and health in the workplace–specifically, reductions in work-related injuries, illnesses, or deaths or reductions in hazardous exposures in the workplace–that can be attributed to NIOSH efforts.

External Factors: Actions or forces beyond NIOSH's control (e.g., by industry, labor, regulators, and other entities) with important bearing on moving research results into practice in the workplace.

SOURCE: Adapted from IOM and NRC (2006).

tp://www.nap.edu/catalog/12639.html

Issues in Developing and Implementing the Evaluation Framework

The variations in the types of NIOSH programs undergoing evaluation raised several issues during the framework committee's work and in the course of the initial set of program evaluations. These issues are further discussed in Chapter 4 and reflect the discussions among NIOSH staff and committee members at the November 2008 workshop held after the evaluations were completed. Primary issues included the following:

- Differences in program mission, including whether a program is dedicated to multiple health and safety issues within a single industry sector (e.g., mining) or focuses on a set of health outcomes (e.g., hearing loss) in whichever industrial sector they occur (a cross-sector program);
- Differences in program structure or management. Programs may use varying management styles and structures, with some programs using a matrix-management approach that coordinates across NIOSH divisions and facilities and others that are located within a single NIOSH division using a more traditional management structure;
- Differences in program size;
- Differences in program and subprogram maturity. An older program or subprogram may have better developed strategic goals and measurable outcomes;
- Overlap among programs;
- The weighting of different activities within and across study areas of a single program;
- The influence of external factors on the implementation, translation, and impact of NIOSH's efforts;
- The definition and difference between intermediate and end outcomes;
- Limitations in the availability of quantitative information (e.g., surveillance data), particularly related to priority setting and documentation of intermediate and end outcomes; and
- The quantitative rating scales. The framework committee was concerned about developing scoring criteria that would maximize consistency across the reports. Many evaluation committee members were initially concerned that the required task of specifying scores for relevance and for impact on 5-point integer

tp://www.nap.edu/catalog/12639.html

scales was too quantitative for a process that, by nature, was largely qualitative.

An additional concern discussed by the framework committee and the evaluation committees was the challenge of evaluating and scoring a federal program within the broader context of agency evaluations being conducted by the Office of Management and Budget using the Program Assessment Rating Tool (PART). Concerns about how the evaluations would be used and their potential impact on the programs' budgets were voiced. NIOSH senior management staff acknowledged the PART evaluation requirements, but also emphasized that NIOSH's reason for conducting thorough and independent external evaluations was focused on quality improvement. The evaluation and framework committees kept both goals in mind and focused on conducting a fair and balanced evaluation.

Revisions to the Evaluation Framework

To monitor and refine the evaluation framework over the course of the evaluations, the framework committee continued to meet once or twice annually from 2006 through 2009 (Appendix A). In 2006 and 2007, the framework committee discussed the first set of ongoing and recently completed evaluations with evaluation committee chairs and members, NIOSH staff, National Academies' staff, and stakeholder groups. They explored the applicability of the framework document, the usefulness of the initial sets of materials provided by NIOSH (the program evidence packages), and issues associated with the quantitative ratings and the ratings criteria. Issues associated with the evaluation of program quality in the evaluation framework were also discussed.

In response, the framework committee released a revised framework in August 2007 to improve clarity, accommodate the revised statement of task as negotiated with NIOSH, and make the document more usable to the evaluation committees. Changes included reorganization of the document to more closely follow the revised statement of task and to improve readability. Slight modifications were made to the criteria for scoring impact and relevance that made the wording more precise. Additions were made in the form of guidance to the evaluation committees regarding scoring a program for relevance based on the committee's ob-

tp://www.nap.edu/catalog/12639.html

served levels of "research priority" and "engagement in appropriate transfer activities."

The four evaluation committees examining the NIOSH Personal Protective Technology, Traumatic Injury, Construction, and Health Hazard Evaluation programs used the August 2007 version of the framework.

As described below, the framework committee continued to receive input from the individual evaluation committees and followed their progress. The framework committee met in 2008 and 2009 with the goals of (1) convening a workshop on lessons learned in the evaluation process, and (2) compiling this report, which includes a further revised framework (version 2009) and recommendations for future evaluations.

Evaluation Committees

As of the publication date of this report, the evaluation of eight NIOSH programs has been completed (IOM and NRC, 2006, 2008, 2009; NRC and IOM, 2007, 2008a,b, 2009a,b). Table 1-1 lists the programs evaluated, and summarizes the activities of the evaluation committees, the version of the framework document applied, and the dates of report release. Framework committee members served as evaluation committee members, liaisons, and report reviewers. Each evaluation committee met three to four times. In addition to meetings, some committees conducted site visits of research program facilities. The NIOSH program provided each committee with an evidence package prior to its first meeting. Additional information was acquired through site visits; input from external stakeholders through open sessions at committee meetings, online questionnaires, and written responses; and oral and written responses to questions posed to the NIOSH programs. Throughout the evaluation process, evaluation committee members provided information to the framework committee on the successes and limitations of the framework in conducting their specific program evaluation.

tp://www.nap.edu/catalog/12639.html

TABLE 1-1 Evaluation Committee Time Lines

Program Name	Evaluation Framework Used	Meeting Dates[a]	Report Release
Hearing Loss	2005	January 5–6, 2006 February 23–24, 2006 March 30–31, 2006	August 2006
Mining Safety and Health	2005	January 12–13, 2006 February 21–22, 2006 May 9–10, 2006	May 2007
Agriculture, Forestry, and Fishing	2005	January 18–20, 2007 March 28–29, 2007 May 30–31, 2007	December 2007
Respiratory Diseases	2005	October 26–27, 2006 December 5–6, 2006 March 22–23, 2007	March 2008
Personal Protective Technology (PPT)	2007[b]	September 27–28, 2007 December 17–18, 2007 March 6–7, 2008	June 2008
Traumatic Injury	2007	March 29–30, 2007 May 31–June 1, 2007 September 6–7, 2007	August 2008
Construction	2007	July 17–18, 2007 September 25–27, 2007 December 10–12, 2007	October 2008
Health Hazard Evaluation (HHE)	2007[b]	October 18–19, 2007 December 10–11, 2007 January 15–16, 2008 February 21–22, 2008	October 2008

[a]The list of meeting dates does not include site visits or committee conference calls.

[b]The PPT and HHE studies used slightly adapted statements of task because of the certification or investigative work of these programs. Therefore the committees had to slightly modify the application of the evaluation framework.

tp://www.nap.edu/catalog/12639.html

Process for Developing this Report

Subsequent to the eight evaluations, the framework committee reconvened for two meetings to discuss the lessons learned during the process and to revise the evaluation framework accordingly. This report is the result of those discussions and revisions. The report draws on the experience of those who participated in developing the framework and those who used the framework to conduct an evaluation. Additionally, the framework committee held a workshop on November 24, 2008, that focused on lessons learned in the evaluation efforts (Appendix A). NIOSH staff members, evaluation committee members and chairs, framework committee members, and IOM and NRC staff members participated in the workshop and shared their perspectives on lessons learned and their ideas for future evaluations. Committee chairs and NIOSH staff members had an additional opportunity for input after the workshop.

NIOSH Follow-Up on the Evaluation Reports

During the time that the framework committee examined the evaluation process, NIOSH staff members have been responding to each of the evaluation reports. NIOSH program staff members were tasked with developing an action plan that addresses the evaluation report findings and recommendations. The action plans are presented to NIOSH senior management and then to the NIOSH Board of Scientific Counselors (BSC). The NIOSH BSC has committed to be actively involved in reviewing the action plans and to write a short assessment of each action plan. Additionally, the BSC is examining the recommendations from all eight of the reports to identify crosscutting issues (e.g., need for improved surveillance). Addressing the issues and strategies identified through this process might provide the greatest efficiency and effectiveness in continuing to improve the relevance and impact of the NIOSH programs.

OVERVIEW OF THIS REPORT

Chapter 2 of this report sets the National Academies' approach to evaluating the NIOSH programs in the larger context of program evaluation. Chapter 3 provides the revised evaluation framework. Because the

tp://www.nap.edu/catalog/12639.html

framework has been integrated into this report (as opposed to previous versions that were stand-alone documents and appendixes to the evaluation reports), a number of the changes involved placing introductory material in Chapter 1. Other changes were more substantive and were the result of the framework committee's careful considerations of the lessons learned, which are detailed in Chapter 4. The concluding chapter provides the committee's recommendations for NIOSH to consider in planning future evaluations. The committee believes the evaluation framework is sufficiently robust and the lessons learned are sufficiently generalizable to offer insights to other federal agencies as they consider program evaluation.

REFERENCES

BLS (Bureau of Labor Statistics). 2008. *Workplace illnesses and injuries in 2007.* http://www.bls.gov/iif/oshwc/osh/os/osnr0030.pdf (accessed April 10, 2009).

BLS. 2009. *Fatal occupational injuries by industry and event or exposure.* http://www.bls.gov/iif/oshwc/cfoi/cftb0223.pdf (accessed April 10, 2009).

IOM and NRC (Institute of Medicine and National Research Council). 2006. *Hearing loss research at NIOSH.* Committee to Review the NIOSH Hearing Loss Research Program. Rpt. No. 1, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

IOM and NRC. 2008. *The personal protective technology program at NIOSH.* Committee to Review the NIOSH Personal Protective Technology Program. Rpt. No. 5, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

IOM and NRC. 2009. *Traumatic injury research at NIOSH.* Committee to Review the NIOSH Traumatic Injury Research Program. Rpt. No. 6, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

Leigh, P. S. 2008. *Cost of occupational injury and illness combining all industries*. Presentation at the November 2008 Seminar for Western Center for Agriculture Health and Safety. http://agcenter.ucdavis.edu/seminar/flyer/2009/Leigh_Nov3_2008.ppt#256,1,Costs of

tp://www.nap.edu/catalog/12639.html

Occupational Injury and Illness Combining All Industries (accessed March 26, 2009).

NIOSH (National Institute for Occupational Safety and Health). 2009a. *NIOSH strategic plan outline, 2004–2009. http://www.cdc.gov/niosh/docs/strategic* (accessed March 19, 2009).

NIOSH. 2009b. *NORA.* http://www.cdc.gov/niosh/nora/ (accessed March 17, 2009).

NIOSH. 2009c. *NIOSH Program Portfolio.* http://www.cdc.gov/niosh/programs/ (accessed March 17, 2009).

NRC and IOM. 2007. *Mining safety and health research at NIOSH.* Committee to Review the NIOSH Mining Safety and Health Research Program. Rpt. No. 2, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2008a. *Agriculture, forestry, and fishing research at NIOSH.* Committee to Review the NIOSH Agriculture, Forestry, and Fishing Research Program. Rpt. No. 3, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2008b. *Respiratory diseases research at NIOSH.* Committee to Review the NIOSH Respiratory Diseases Research Program. Rpt. No. 4, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2009a. *The health hazard evaluation program at NIOSH.* Committee to Review the NIOSH Health Hazard Evaluation Program. Rpt. No. 7, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2009b. *Construction research at NIOSH.* Committee to Review the NIOSH Construction Research Program. Rpt. No. 8, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

tp://www.nap.edu/catalog/12639.html

2

The Program Evaluation Context[1]

Program evaluation has been defined as "systematic inquiry that describes and explains the policies' and program's operations, effects, justifications, and social implications" (Mark et al., 2000, p. 3) or " . . . the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve program effectiveness, and/or inform decisions about future programming"(Patton, 1997, p. 23). The evaluations of National Institute for Occupational Safety and Health (NIOSH) programs carried out under the guidance of the framework committee represent just one way of evaluating research programs. This chapter places the National Academies' evaluations of NIOSH programs in context by providing a brief overview of the general principles involved in program evaluation and by describing where the process fits in the spectrum of current practices in evaluating research programs. At the conclusion of some of the overview sections throughout the chapter, the committee's findings specific to the evaluation process used by the framework and evaluation committees are included in bold and italicized text.

PROGRAM EVALUATION

Although formal program evaluations, especially of educational programs, preceded World War I, the profession as currently practiced in the United States has increased in prominence within the past 50 years. A

[1]This chapter draws on background papers commissioned by the committee from Sonia Gatchair, Georgia Institute of Technology, and Monica Gaughan, University of Georgia.

PREPUBLICATION COPY—UNCORRECTED PROOFS

tp://www.nap.edu/catalog/12639.html

major impetus to this growth was the need to assess the social programs instituted through the War on Poverty and Great Society policies of the 1960s (Shadish et al., 1991). Legislative requirements for the evaluation of many programs represented a turning point in the growth in the number of evaluations. Evaluation is now an established professional practice, reflected through organizations such as the American Evaluation Association and the European Evaluation Society (AEA, 2009; EES, 2009). Program evaluation is one element of results-oriented management, the approach to public management embodied in the past decade in the Government Performance and Results Act (OMB, 2009a) and the Office of Management and Budget's (OMB's) Program Assessment Rating Tool (OMB, 2009b).

Current efforts in program evaluation follow several schools of thought that differ in the evaluation processes used but are all focused on achieving a valid evaluation. The essence of evaluation is determining what is of value in a program. The work revolves around understanding program goals (if available), setting criteria for success, and gathering information to determine whether the criteria are being met as a result of program activities. Program evaluations focus on examining the characteristics of a portfolio of projects rather than assessing one project at a time and often use retrospective information about program outputs and outcomes. Program evaluation differs from a research project in being more tightly connected to practice; it is commissioned by a specific user or organization and designed to inform decision making. It also differs from performance measurement, which is an ongoing process that gathers indicators of what the program is accomplishing but may not assess why the indicators are changing.

Program evaluations can serve several functions. When the program is initially in development or is undergoing changes and is being evaluated with the goal of program improvement, the evaluation is termed a *formative evaluation* (Scriven, 1991). These evaluations are often initiated and used in-house. When the objective of the evaluation is to assess the program's outcomes in order to determine whether the program is succeeding or has accomplished its goals, the evaluation is termed a *summative evaluation* (Scriven, 1967; Gredler, 1996). Users of summative evaluations are often decision makers outside of the program. Program evaluation often also helps communicate the program's goals and accomplishments to external audiences. Evaluations provide information that contributes to decisions that shape program goals, strategic plans, and actions. In these cases, they serve instrumental functions. Often they

tp://www.nap.edu/catalog/12639.html

also serve enlightenment functions, such as increasing general understanding of program operations, underlying assumptions, or social context (Weiss, 1977).

The practice of evaluating research programs has historically has been somewhat separate from that of social program evaluation. Qualitative assessments of research programs in the United States date back to the 1950s (NAS, 1959). The evaluation of research programs took a more quantitative turn in the 1970s as evaluations started to draw on the new availability of large-scale databases to describe scientific activity. Research program evaluation is distinguished from social program evaluation in a number of ways, including the dominant use of peer-review panels and the use of specialized data, including publication and patent-based measures (see discussion later in this chapter).

The evaluations of NIOSH programs discussed in this report were undertaken in the context of the externally mandated Program Assessment Rating Tool process, a summative evaluation process developed by OMB. However, NIOSH leadership established their primary goal as program improvement, making the evaluations primarily formative.

LOGIC MODELS

The evaluations of NIOSH programs used logic models—both a general logic model for NIOSH research and specific logic models for each program evaluated. Prior to the work of the evaluation committees, NIOSH contracted with RAND Corporation to provide operational and analytical assistance with compiling the evidence packages for the reviews and developing the logic models; a detailed description of that effort can be found in a recent RAND report (Williams et al., 2009).

Logic models are widely used in program evaluation (W. K. Kellogg Foundation, 2000; World Bank, 2000) to represent visually what evaluators call "program theory." This phrase refers to the understanding of how the program is supposed to work. How do the program resources become results, and through what channels do those results have their expected impacts? The logic model may be represented as a set of boxes and arrows or as a hierarchy of goals, intermediate outcomes, and final outcomes. The representation provides guidance for the evaluation by

tp://www.nap.edu/catalog/12639.html

pointing to relevant kinds of information to be considered in the assessment and often to indicators in the various areas of the model.

McLaughlin and Jordan (1999) refer to logic models as a way of "telling your program's performance story." The common elements of logic models are inputs, activities, outputs, customers, and outcomes (short, medium, and long term), plus external influences (Wholey, 1983; Figure 2-1).

Building a logic model is a process that should involve a team of people with different roles in the program who interact with external stakeholders at many points. After collecting relevant information and clearly identifying the problem the program addresses, the team organizes its information into various elements and composes a diagram that "captures the logical flow and linkages that exist in any performance story" (McLaughlin and Jordan, 1999, p. 68).

Logic models are nearly always predominantly linear and causal because agencies use them to think through how programs will achieve their public goals. In research planning and evaluation, this linearity is ironic. The widespread consensus is that research does not create its effects in a linear fashion. Rather, it is embedded in a complex ecology of relationships that shape and spread knowledge through a variety of channels, not just research knowledge.

Additionally, it is challenging for logic models to capture some outputs such as the development of human capital. Over time, a program may have a significant impact on a field by helping to build a community of practitioners and researchers. For example, NIOSH's impact on the existence and growth of the occupational safety and health research community is hard to capture in a logic model. In addition, ongoing dialogues with external stakeholders shape research activities and spread research knowledge in ways that are hard to track.

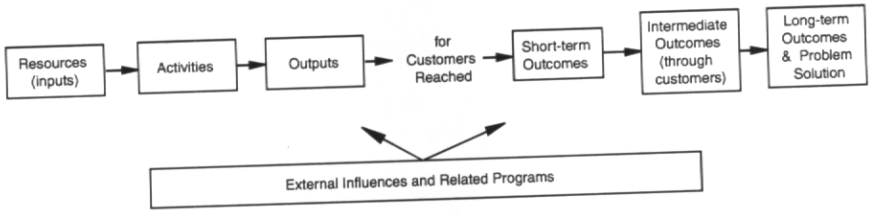

FIGURE 2-1 Elements of the logic model. Reprinted from McLaughlin and Jordan, 1999 with permission from Elsevier.

tp://www.nap.edu/catalog/12639.html

Program evaluations that solely rely on the logic model almost inevitably miss information on some of the nonlinear effects of program activities.

> ***The logic models used in the evaluation of NIOSH programs helped program staff and evaluators organize information into steps in the flow of program logic. However, because some of the NIOSH programs spanned several NIOSH divisions and laboratories, the logic model sometimes made it hard for the evaluation committee to grasp the full picture of the research program. Furthermore, the logic models focused a great deal of attention on the most readily observable short- and medium-term outcomes, perhaps missing information on nonlinear and more diffuse contributions of the programs to the development of knowledge and human capital in occupational safety and health.***

ROLE OF STAKEHOLDERS

The practice of program evaluation has paid special attention to external stakeholders and the role they play in the evaluation process. Sometimes stakeholders are direct beneficiaries of the program; for example, for a day-care center program, the major stakeholders are the families whose children receive care. Sometimes the stakeholders are organizations with whom the program must work to achieve its goals. In the case of research on occupational safety and health, key stakeholders include workers, employers, and regulatory agencies.

Stakeholder participation in evaluating research programs has come more slowly than in social program evaluation. Early evaluation panels tended to consist entirely of scientists and engineers. But as research policy became more focused on making research relevant to the private sector, evaluation panels also began to include industry and labor representation, often scientists and engineers working in industry and labor organizations. Individuals and families exposed to environmental hazards often organize to increase research and remediation efforts, and stakeholders from these groups also participate in evaluation processes.

Just as social program evaluation pays particular attention to differences in knowledge and expertise between evaluators and stakeholders, in the evaluation of research programs the different contributions of sci-

tp://www.nap.edu/catalog/12639.html

entific experts and external stakeholders both need to be respected. When the research being evaluated is intended to serve vulnerable populations, current standard practice in the evaluation of research programs, as described in the last paragraph, is not sufficient to give voice to these groups and additional attention needs to be paid to obtaining adequate input.

The National Academies evaluation committees included a variety of members with strong connections to NIOSH's external stakeholder groups, such as manufacturers of safety equipment, labor organizations, and employers. The committees also reached out to a wide range of external stakeholder groups for input, including vulnerable worker populations.

METHODS OF EVALUATION

Evaluations of research programs necessarily use a variety of approaches. Expert panel review is the "bread-and-butter" approach worldwide, but there is also a long track record of evaluation studies, in which external consultants gather and analyze primary data to inform the expert deliberations.

Within the range of evaluation approaches for research programs, the National Academies' evaluations of NIOSH programs clearly fall among expert panel evaluations, rather than evaluation studies.

Expert Panel Review

Merit review, peer review, and expert panels are used widely for both ex ante and ex post evaluations of the productivity, quality, and impact of funding organizations, research programs, and scientific activity. Benefits and limitations of this approach have been reviewed extensively (Bozeman, 1993; Guston, 2003; Hackett and Chubin, 2003). Expert panel review is the oldest—and still most common—form of research and development evaluation. In fact, the expert panel is very much a historical development from the National Academies itself, which was es-

tp://www.nap.edu/catalog/12639.html

tablished in the 19th century to provide scientific and technical policy advice to the federal government. The underlying evaluative theory of the expert panel is that scientists are uniquely able to evaluate the quality and importance of scientific research (Kostoff, 1997). The preeminence of scientists to evaluate the quality and importance of scientific research was further codified in the research agencies that developed under the philosophy of Vannevar Bush in the 1940s (Bush, 1945).

Expert judgment is particularly capable of evaluating the quality of discrete scientific research activities and the relevance of such discrete activities to particular bodies of knowledge. For example, toxicologists and biochemists—through their scientific training—are uniquely capable of assessing the contributions of particular theories, research methodologies, and evidence to answer specific scientific questions and problems. The major limitation of expert panel review is that traditional training and experience in the natural and physical sciences do not prepare scientists to address questions related to the management, effectiveness, and impact of the types of broad research portfolios that federal agencies typically manage.

Although expert panel reviews work to balance conflicting values, objectives, or viewpoints, they also may lead to tensions in the same areas they are expected to resolve. As noted above, the review process may be broadened to include other stakeholders beyond "experts" or "peers." Expert panels usually operate with an evaluation protocol developed by an outside group, including evaluation procedures, questions to be answered, and evaluation criteria (e.g., the evaluation of the Sea Grant College Program, Box 2-1). The panels usually review a compilation of data on the program, including plans, input counts (budget, staffing), project descriptions, and lists of results. They then talk with individuals connected to the program, both inside and outside the given agency, and use their own experience and judgment in reaching conclusions.

Closely tied to review processes is the assignment of various types of ratings. For example, the Research Assessment Exercise of the United Kingdom uses 15 panels and 67 subpanels following a common protocol to assess university research programs and assign scores by discipline area (RAE, 2009). Rating scales are being used more frequently as

tp://www.nap.edu/catalog/12639.html

BOX 2-1
Evaluation of the National Sea Grant College Program

The National Sea Grant College Program, funded by the National Oceanic and Atmospheric Administration, is a nationwide network of 30 university programs aimed at conducting research, education, and training on coastal resources and marine policy. A 1994 National Academies' review of the program (NRC, 1994) recommended that individual program evaluations be conducted on a 4-year review cycle. From 1998 to 2006, two cycles of site visit evaluations were conducted using a uniform and detailed set of performance criteria and a standardized set of benchmarks and indicators developed by the external review panel charged with oversight (NRC, 2006). Programs were scored on criteria in the major areas of:

- Using effective and aggressive long-range planning;
- Organizing and managing for success;
- Connecting Sea Grant with users; and
- Producing significant results.

At the end of the 4-year cycle, a final evaluation process provided a comparative assessment across the 30 university programs.

The National Academies was asked to examine the National Sea Grant evaluation process. The resulting report included recommendations emphasizing the need for internal assessments to complement external evaluations, increased opportunities for interactions among the university programs, streamlined annual assessments, and improvements in strategic planning (NRC, 2006).

evaluations have become more and more oriented to demonstrating performance to outside audiences or to allocating resources. Rating scales capture qualitative judgments on ordinal scales and allow for descriptions of performance at the various levels.

Characteristics that are sought in expert panel reviews include a panel with a balanced set of expertise and credibility among various stakeholder groups and independence and avoidance of conflict of interest among panel members to the extent possible. Selection of panel members can involve trade-offs between recruiting independent reviewers or recruiting reviewers with knowledge and understanding of the program and its field of science. For this reason, expert review panels are seldom completely free of bias and may have conflicts of interest; the preferred practice, of course, is for conflicts to be considered and disclosed. Independence is also reinforced when the panel is commissioned by, and reports to, a user located at least one level above the program in

tp://www.nap.edu/catalog/12639.html

the management hierarchy. The panel adds value by including its perspectives and insights in its report. The panel makes the evidence base for its conclusions explicit in the report and usually makes a limited number of realistic recommendations, phrased broadly enough to allow management to adapt the recommendations to specific circumstances.

The National Academies committees follow a thorough bias and conflict-of-interest process that includes completion of disclosure forms and the bias and conflict-of-interest discussion held at the first meeting.

Other Methods of Evaluating Research Programs

Other types of evaluations generally involve hiring consultants to provide analyses of specific outputs of the program. Because the goal of a research program is new knowledge, publications represent a concrete and observable manifestation of new knowledge and are frequently used as a convenient measure of research program outputs. Publications in peer-reviewed journals provide an indication of quality control, and citations to published articles are used to assess the scientific impact of the work. Patents provide a similar set of measures for technology development. Thus, evaluations of research programs have extensive relevant datasets on which to base their assessments.

Statistical analyses of data on publications (e.g., books, journal articles, review articles, book chapters, notes, letters) range from fairly simple counts and comparisons of publications to highly sophisticated factor analyses and correlations of many types of terms, such as keywords, institutions, and addresses, that lead to the development of networks or maps of the ways in which the research outputs are connected. These bibliometric methods are used extensively to evaluate research activities and compare research output across institutions, disciplines, fields, funding programs, countries, and groups of researchers (Kostoff, 1995; Georghiou and Roessner, 2000; Hicks et al., 2004; Weingart, 2005). Bibliometric methods also can be used to assess the extent of collaboration. Visualization techniques now produce "maps of science" allowing organizations that support research to "see" where the work they have supported fits into research in a specific field or the extent to which it is being used in other research endeavors. An important strength of bibliometric analyses is that they are data-based analyses following a fixed

tp://www.nap.edu/catalog/12639.html

set of rules or algorithms. The analyses are often used as a complement to peer-review techniques, surveys, or impact analyses of research activities. An important weakness, however, is that the measures are incomplete. They do not capture all the dimensions of performance or its context, factors that an evaluation usually needs to consider. In general, a composite set of measures is used to determine the effectiveness of the research activities, institutions, or national programs (Box 2-2).

Other methods used in evaluating research programs include methodologies drawn from the social sciences, including case studies, interviews, and surveys. One special application of case studies in the evaluation of a research program, for example, is the TRACES approach, named for an early study of Technology in Retrospect and Critical Events in Science (IIT, 1968). This approach starts from a recent accomplishment or success, then tracks the complex set of earlier research results and technologies that made it possible. Programs with economic goals have also used case studies to illustrate the return on investment in advanced technology projects (Ruegg, 2006).

BOX 2-2
Review of the National Science Foundation's Science and Technology Center Programs

Beginning in 1989, the National Science Foundation (NSF) established 25 Science and Technology Centers (STCs) across the United States. The goal was to promote cutting-edge fundamental research in all areas of science, improve the quality of science and math education, and enhance the transfer of knowledge among disciplines. The efforts of these center programs have been evaluated through several external assessments, including site-visit teams. A congressionally requested review of the management of the STC program was conducted by the National Academy of Public Administration (NAPA, 1995).

The National Academies was asked to conduct an evaluation of the accomplishments of the STC Program as a whole, rather than individual center evaluations (NRC, 1996). Evaluation input included data from Abt Associates regarding their historical review; secondary data analysis on the characteristics and operations of the 25 centers; bibliometric and patent analyses; and surveys of principal investigators, industry/federal laboratory representatives, educational outreach collaborators, and other key stakeholders (Fitzsimmons et al., 1996). The National Research Council report recommendations included an increased emphasis on graduate and undergraduate education and coordination of the reviews of the program (NRC, 1996).

tp://www.nap.edu/catalog/12639.html

SUMMARY

In summary, the evaluation of research programs is an established branch of program evaluation. The National Academies' evaluation of NIOSH research programs used one of the most common approaches: expert panel review. As is common in evaluations of applied research programs, this process involved stakeholders as members of the evaluation committees and also sought external stakeholder input. The evaluation framework described in Chapter 3 organizes data into a common evaluation tool based on a logic model approach and provides for consideration of external factors. Similar to many research program evaluation efforts, the evaluation committees used this structured rating tool to provide some consistency in ratings across programs. The process did not, however, expand into an evaluation study by gathering new data or extensively analyzing external data sources. The evaluations of NIOSH programs fall well within the range of acceptable practice in evaluating research programs and are compiled in comprehensive reports that went through peer review under the National Academies' report review process.

REFERENCES

AEA (American Evaluation Association). 2009. *American Evaluation Association.* http://www.eval.org (accessed March 23, 2009).

Bozeman, B. 1993. Peer review and evaluation of R&D impacts. In *Evaluating R&D impacts: Methods and practice.* Edited by B. Bozeman and J. Melkers. Boston, MA: Kluwer Academic. Pp. 79–98.

Bush, V. 1945. *Science: The endless frontier.* Washington, DC: U.S. Government Printing Office.

EES (European Evaluation Society). 2009. *European Evaluation Society.* http://www.europeanevaluation.org/ (accessed March 23, 2009).

Fitzsimmons S. J., O. Grad, and B. Lal. 1996. *An evaluation of the NSF Science Technology Centers Program.* Vol. 1, Summary. Washington, DC: Abt Associates. http://www.nsf.gov/od/oia/programs/stc/reports/abt.pdf (accessed March 23, 2009).

Georghiou, L., and D. Roessner. 2000. Evaluating technology programs: Tools and methods. *Research Policy* 29(4–5):657–678.

Gredler, M.E. 1996. *Program evaluation.* Englewood Cliffs, NJ: Merrill.

tp://www.nap.edu/catalog/12639.html

Guston, D. 2003. The expanding role of peer review processes in the United States. In *Learning from science and technology policy evaluation: Experiences from the United States and Europe.* Edited by P. Shapira and S. Kuhlmann. Cheltenham, UK, and Northampton, MA: Edward Elgar. Pp. 81–97.

Hackett, E., and D. Chubin. 2003. *Peer review for the 21st century: Applications to education research.* Paper presented at the National Research Council Workshop, Washington, D.C., February 25, 2003. http://www7.nationalacademies.org/core/HacketChubin_peer_review_paper.pdf (accessed November 16, 2008).

Hicks, D., H. Tomizawa, Y. Saitoh, and S. Kobayashi. 2004. Bibliometric techniques in the evaluation of federally funded research in the United States. *Research Evaluation* 13(2):78–86.

IIT (Illinois Institute of Technology). 1968. *Technology in retrospect and critical events in science.* Vol. 1. Chicago, IL: IIT Research Institute.

Kostoff, R. N. 1995. Federal research impact assessment—axioms, approaches, applications. *Scientometrics* 34(2):163–206.

Kostoff, R. N. 1997. Peer review: The appropriate GPRA metric for research. *Science* 277:651–652.

Mark, M., G. Henry, and G. Julnes. 2000. *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs.* San Francisco: Jossey-Bass.

McLaughlin, J. A., and G. B. Jordan. 1999. Logic models: A tool for telling your program's performance story. *Evaluation and Program Planning* 22:65–72.

NAPA (National Academy of Public Administration). 1995. *National Science Foundation's Science and Technology Centers: Building an interdisciplinary research program.* Washington, DC: NAPA.

NAS (National Academy of Sciences). 1959. *Panel reports of the NAS-NRC panels advisory to the National Bureau of Standards.* Washington, DC: National Academy of Sciences.

NRC (National Research Council). 1994. *A review of the NOAA National Sea Grant College Program.* Washington, DC: National Academy Press.

NRC. 1996. *An assessment of the National Science Foundation's Science and Technology Centers Program.* Washington, DC: National Academy Press.

NRC. 2006. *Evaluation of the Sea Grant Program review process.* Washington, DC: The National Academies Press.

tp://www.nap.edu/catalog/12639.html

OMB (Office of Management and Budget). 2009a. *Government Performance and Results Act.* http://www.whitehouse.gov/omb/mgmt-gpra/gplaw2m.html (accessed March 20, 2009).

OMB. 2009b. *Assessing program performance.* http://www.whitehouse.gov/omb/part/ (accessed March 20, 2009).

Patton, M.Q. 1997. *Utilization-focused evaluation: The new century text.* Thousand Oaks, CA: Sage.

RAE (Research Assessment Exercise). 2009. *Research Assessment Exercise.* http://www.rae.ac.uk/ (accessed March 23, 2009).

Ruegg, R. 2006. *Bridging from project case study to portfolio analysis in a public R&D program—A framework for evaluation and introduction.* NIST GCR 06-891. http://www.atp.nist.gov/eao/gcr06-891/gcr06-891report.pdf (accessed March 23, 2009).

Scriven, M. 1967. The methodology of evaluation. In *Curriculum evaluation: AERA monograph series on evaluation.* Edited by R. E. Stake. Chicago: Rand McNally. Pp. 39–85.

Scriven, M. 1991. Beyond formative and summative evaluation. In *Evaluation and education: At quarter century.* Edited by M. W. McLaughlin and D. C. Phillips. Chicago: University of Chicago Press. Pp. 19–64.

Shadish, W.R., T. D. Cook, and L. C. Leviton. 1991. *Foundations of program evaluation: Theories of practice.* Newbury Park, CA: Sage.

W. K. Kellogg Foundation. 2000. *Logic model development guide.* http://www.wkkf.org/Pubs/Tools/Evaluation/Pub3669.pdf (accessed March 23, 2009).

Weingart, P. 2005. Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics* 62(1):117–131.

Weiss, C. H. 1977. Research for policy's sake: The enlightenment function of social research. *Policy Analysis*: 3:531–545.

Wholey, J. S. 1983. *Evaluation and effective public management.* Boston: Little, Brown.

Williams, V. L., E. Eiseman, E. Landree, and D. M. Adamson. 2009. *Demonstrating and communicating research impact: Preparing NIOSH programs for external review*. Santa Monica, CA: RAND.

World Bank. 2000. *Logframe handbook: A logical framework approach to project cycle management.* http://www.wau.boku.ac.at/fileadmin/_/H81/H811/Skripten/811332/811332_G3_log-framehandbook.pdf (accessed March 23, 2009).

tp://www.nap.edu/catalog/12639.html

tp://www.nap.edu/catalog/12639.html

3

Evaluation Framework

As discussed throughout this report, an evaluation framework was developed to guide the NIOSH program evaluations and provide a set of criteria to be used in scoring the relevance and impact of each program's efforts in reducing work-related hazardous exposures, illnesses, and injuries. The National Academies Committee for the Review of NIOSH Research Programs (referred to as the framework committee) developed an initial evaluation framework that was honed and refined to improve its utility, clarity, and emphasis as the eight evaluations proceeded. The framework presented in this chapter differs slightly from the versions that were used by the evaluation committees.[1] Insights gained in the evaluation of the eight NIOSH programs (Chapter 4) are included in this version of the framework, which is provided for consideration in future program evaluations. In addition to evaluating NIOSH programs, this analytic framework and approach may be applicable to the evaluation of other federal agency research programs or research programs in other organizations.

In conducting its evaluation, each committee was asked to determine whether the NIOSH program was undertaking high-priority, relevant research and transfer activities (relevance) and whether these efforts are improving health and safety in the workplace (impact). The evaluation committee was also tasked with (1) rating both the relevance and the impact of the NIOSH program using 1–to–5 integer scales, and (2) providing

[1]The evaluation framework document used by the individual evaluation committee is provided as an appendix in each of the evaluation committee reports (IOM and NRC, 2006, 2008, 2009; NRC and IOM, 2007, 2008a,b, 2009a,b).

PREPUBLICATION COPY—UNCORRECTED PROOFS

tp://www.nap.edu/catalog/12639.html

input about emerging areas of research and recommendations for program improvement.

OVERVIEW OF THE EVALUATION FRAMEWORK

After examining different approaches to program evaluation (Chapter 2), the framework committee decided to define the scope and stages of the evaluation process based on the logic model (Williams et al., 2009). The resulting evaluation framework described in this chapter breaks the logic models developed by NIOSH (Figure 3-1) into discrete program components to be assessed by each evaluation committee. Criteria for evaluation of each component of the framework are detailed below. In the evaluation framework (overview provided in Figure 3-2), the assessment of strategic goals and objectives, inputs, activities, and outputs (B to E) largely define the relevance of the program, while assessment of intermediate and end outcomes (F and G) largely define the program impact.

The following major components of each NIOSH program were assessed by the evaluation committees:

- Major occupational safety and health *challenges* in the program area.
- *Goals and objectives* as defined by NIOSH.
- *Inputs* (e.g., budget; staff; facilities; and input from the program's research management, the NIOSH Board of Scientific Counselors, and stakeholders).
- *Activities* (efforts by NIOSH staff, contractors, and grantees; e.g., surveillance of injury, illness, and hazards; exposure assessment research; health-effects research; injury-risk factor research; intervention research; health services research; and technology transfer activities).
- *Outputs* (NIOSH products; e.g., publications, reports, conferences, databases, tools, methods, guidelines, recommendations, education and training, and patents).
- *Intermediate outcomes* (actions by external stakeholders in response to NIOSH products; e.g., policy change, training and education, self-reported use or repackaging of NIOSH data by

tp://www.nap.edu/catalog/12639.html

stakeholders, adoption of NIOSH-developed technologies, implemented guidelines, and licenses).

- *End outcomes* (e.g., reduction in work-related injuries, illnesses, or hazardous exposures in the workplace).

The framework committee understood that the efforts of any research program or the evaluation of that program will not be as linear as presented in either Figure 3-2 or Box 3-1; rather, they are iterative processes. Overlap necessarily occurs between the assessment of relevance and impact, particularly in the assessment of information transfer. Furthermore, components of any program may not fit perfectly into any one category. For example, training and development programs are appropriately defined as outputs by NIOSH in the logic model (Figure 3-1), but the framework committee found more value in focusing on the responses to these outputs as intermediate outcomes in the evaluation.

Some NIOSH programs are organized using a matrix management approach as they span several NIOSH divisions or laboratories. Because resources within NIOSH are allocated in large part at the division level, rather than the program level, a matrix organization may have little control over the input portion of the logic model and therefore fewer resources within its direct control on which to base decisions.

Following the suggested evaluation process ensured a level of consistency and comparability among all the evaluation committees. For future program evaluations, training on logic models and criteria for differentiating the various components of the logic model would be beneficial at the inception of the evaluation process both for NIOSH staff as they assemble the evidence packages and for evaluation committee members as they begin their assessment of the program.

http://www.nap.edu/catalog/12639.html

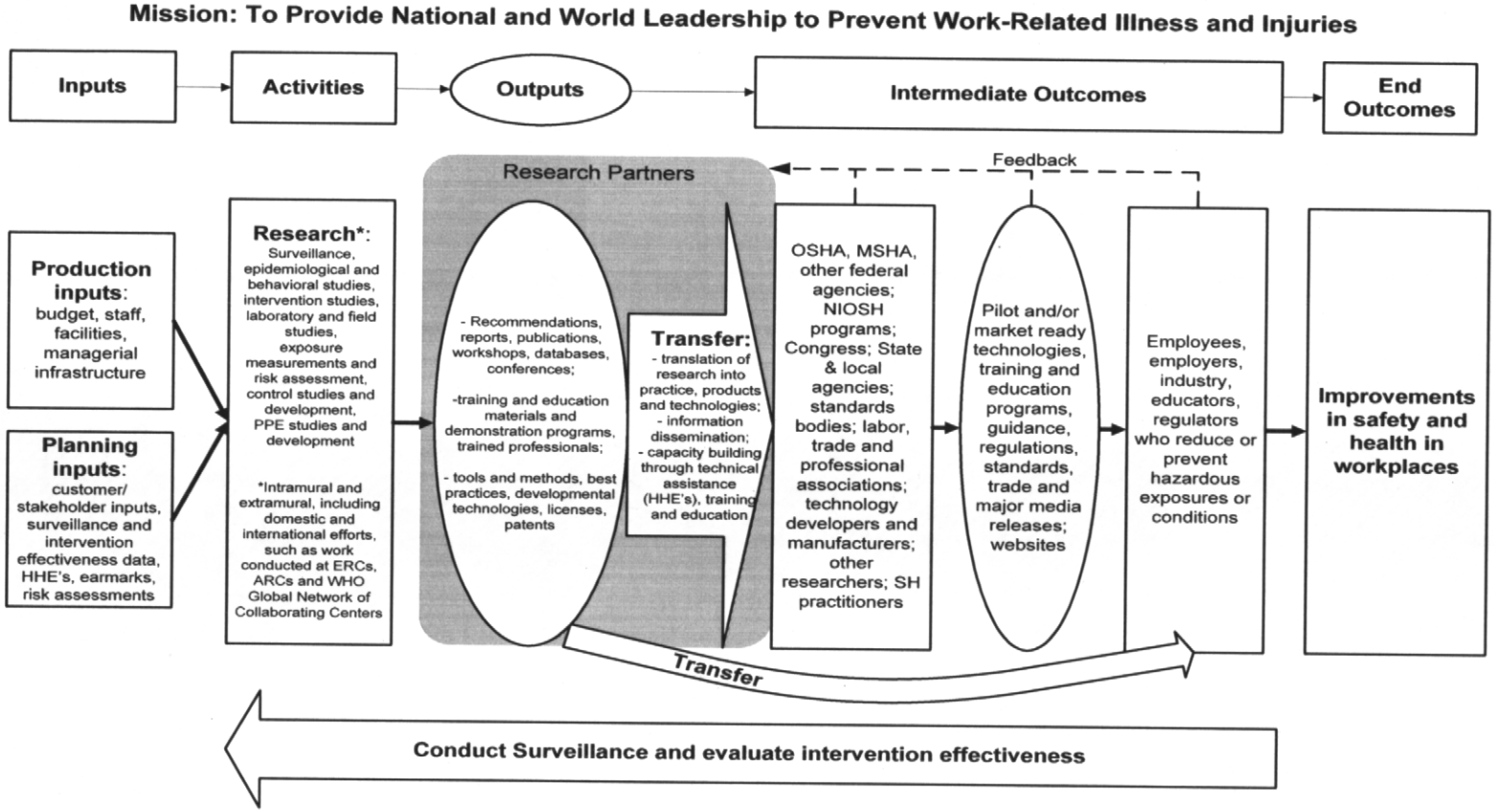

FIGURE 3-1 The National Institute for Occupational Safety and Health logic model.

tp://www.nap.edu/catalog/12639.html

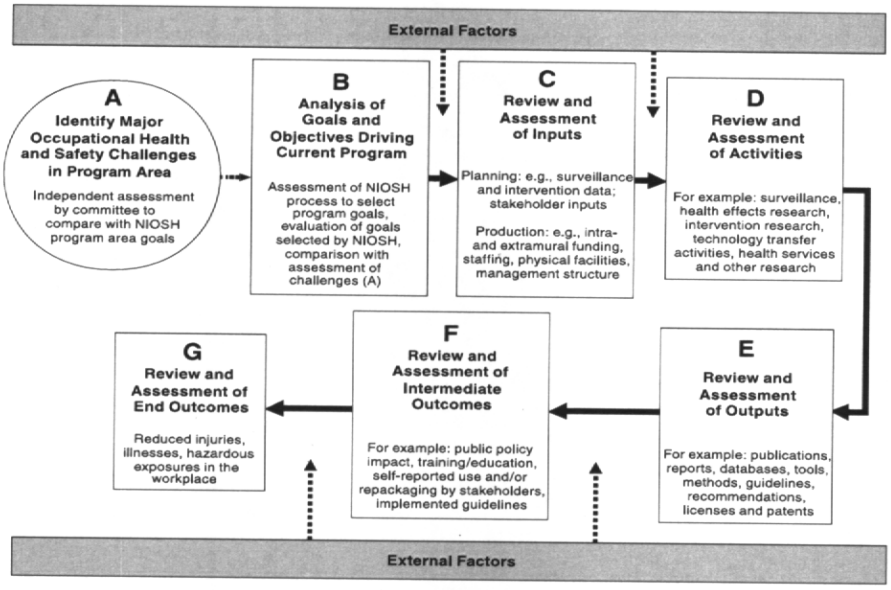

FIGURE 3-2 Overview of the evaluation process.
NOTE: EC = evaluation committee

Drawing on the program logic model, the evaluation framework, and the evaluation committee members' expertise, the evaluation committees began by examining important inputs and external factors affecting the NIOSH research program's agenda. Examples of external factors included research activities of industry and other federal agencies as well as the political and regulatory environment. The evaluation then focused on the program's research activities, outputs, associated transfer activities, and resulting intermediate and end outcomes.

Box 3-1 provides a summary of the evaluation process as suggested by the framework committee. Detailed guidance on each step is provided in later sections of this chapter. The following key factors were considered in assessing the relevance of NIOSH research programs:

- The severity and/or frequency of health and safety hazards addressed and the number of people at risk (magnitude) for these hazards.
- The extent to which NIOSH research programs identified and addressed gender-related issues and issues of vulnerable popula-

tp://www.nap.edu/catalog/12639.html

tions[2] and the extent to which NIOSH research programs addressed the health and safety needs of small businesses.

- The stage of research on the problems being addressed. As the health effects are understood, research efforts should shift from etiologic research to intervention research and then to intervention-effectiveness research. Gaps in the spectrum of prevention need to be addressed; for example, research on exposure assessment may be necessary before the next intervention steps can be taken.
- The structure, in addition to the content, of the research program. A relevant research program is more than a set of unrelated research projects; it is an integrated program involving interrelated surveillance, research, and transfer activities.
- Appropriate NIOSH consideration of external stakeholder input.

The evaluation committees had the option to consider these and other relevant factors as they progressed through each stage of the evaluation.

BOX 3-1
Steps in the Evaluation Process

1. Gather appropriate information.
2. Assess external factors.
3. Identify time frame to be evaluated.
4. Identify major occupational health and safety challenges in program area.
5. Analyze program goals and objectives.
6. Identify major program components.
7. Evaluate program inputs, activities, outputs, and outcomes
8. Determine scores for relevance and impact and provide the rationale.
9. Assess the NIOSH process for targeting priority research needs and provide the committee's assessment of emerging issues.
10. Prepare report by using the template provided as a guide.

[2]Vulnerable populations are defined as groups of workers who have biological, social, or economic characteristics that place them at increased risk for work-related conditions or on whom inadequate data have been collected. These populations include low-wage workers, disadvantaged minorities, disabled persons, and non-English-speakers for whom language or other barriers present health or safety risks.

tp://www.nap.edu/catalog/12639.html

Data documenting end outcomes are often quite limited or are not available to quantify reductions in illness, injury, and hazardous exposures. Data documenting intermediate outcomes, although likely also limited, could serve as an appropriate proxy for end outcome data if the relationship between occupational exposures and health outcomes is well understood. For example, changes in regulations or procedures likely to result in reduced exposures are important measures of intermediate outcomes.

Useful program evaluation requires specific questions and criteria for assessing each component of the program; a disciplined focus on a small number of questions or hypotheses typically related to program goals, performance criteria, and performance standards; a rigorous method of answering the questions or testing the hypotheses; and a credible procedure for developing qualitative and quantitative assessments. Because of the uniqueness of each NIOSH program, each evaluation committee determined the most reasonable way to apply the evaluation criteria.

EVALUATION COMMITTEES

The individual evaluation committees were formed in accordance with the rules of the National Academies that focus on ensuring a balanced committee. Each evaluation committee included persons with expertise appropriate for the specific NIOSH research program under review (including researchers and representatives of stakeholder groups; e.g., worker organizations and industry), experts in technology and knowledge transfer, and experts in program evaluation.

The evaluation committee gathered appropriate information from the NIOSH research program under review, external stakeholders affected directly by the NIOSH program research, and relevant independent parties. The original contracts between NIOSH and the National Academies specified that each evaluation committee would consist of about 10 members, meet three times, and prepare a report due to NIOSH within 9 months of the first meeting of the evaluation committee. As noted in Chapter 4, future evaluations may consider extending the time frame to 12 to 14 months, depending on the size and complexity of the program being evaluated.

tp://www.nap.edu/catalog/12639.html

STEPS IN THE EVALUATION PROCESS

The evaluation process consists of 10 steps, described in the following sections and summarized in Box 3-1. A description is provided of how NIOSH programs were evaluated by the National Academies. This model can be applied to other program evaluations as well.

1. Gather Appropriate Information

Each NIOSH program under review provided information to the relevant evaluation committee, including that outlined in Box 3-2. Some of the evaluation committees requested additional information from the program. Organizing the information listed in Box 3-2 by goal (or by subprogram if organizing by goal was not feasible) was helpful to the committees.

In addition to the information provided by the NIOSH program, the evaluation committees independently collected additional information as deemed necessary for the evaluation, such as the perspectives of external stakeholders, including the Occupational Safety and Health Administration (OSHA), the Mine Safety and Health Administration (MSHA), workforces and their unions, and industry. In conducting the review, the evaluation committees examined how the inputs, activities, outputs, and intermediate outcomes contribute to the impact and relevance of the program as a whole.

BOX 3-2
Evaluation Committee Information Needs

- Program background and overview:
  - Program history
  - Program management structure
  - Major program challenges
  - Program goals and objectives, past (for period under review) and current
  - Process for developing and updating program strategic plans
  - Enabling or authorizing legislation
  - Major subprograms (if appropriate)
  - Results of previous program reviews (e.g., annual review by NIOSH leadership team or external scientific program reviews)
  - External factors affecting the program

tp://www.nap.edu/catalog/12639.html

- Interactions with external stakeholders and with other NIOSH programs:
  - The role of program research staff in NIOSH policy setting, OSHA and MSHA standard setting, voluntary standard setting, and other government policy functions
  - Interactions and working relationships with other NIOSH programs
  - Identification of other institutions and research programs with similar portfolios and an explanation of the relationship between NIOSH activities and those of other institutions
  - Key partnerships with employers, labor, other government organizations, academic institutions, nonprofit organizations, and international organizations

- Program inputs:
  - Production inputs (program resources):
    - Funding by year for period under review
    - Funding by objective or subprogram
    - Program staffing, full-time equivalents, and laboratory facilities, by subprogram (if indicated)
    - Percentage of program budget that is discretionary (beyond salaries)
    - Percentage of program budget that is earmarked
    - Significant contributions to the program from other sources (in kind or funds)
  - Planning inputs:
    - Surveillance data, inputs from the Health Hazard Evaluation and the Fatality Assessment and Control Evaluation programs, and intramural and extramural research findings that influenced program goals and objectives
    - Planning inputs from stakeholders such as advisory groups; National Occupational Research Agenda (NORA) teams; and professional, industry, and labor groups (specify if any input comes from groups representing small business or vulnerable populations)
    - Related OSHA and MSHA strategic plans or other input
    - Process for soliciting and approving intramural research ideas
    - Process for soliciting and approving program-supported extramural research activities

- Program activities (more details provided in Box 3-3):
  - Intramural:
    - Surveillance activities
    - Research activities
    - Transfer activities to encourage implementation of research results for improved occupational safety and health (e.g., information dissemination, technical assistance, and technology and knowledge transfer)
    - Key collaborations in intramural activities (e.g., with other

tp://www.nap.edu/catalog/12639.html

government agencies, academe, industry, and unions)
    - Extramural funded by NIOSH:
        - Requests for applications developed by program
        - Funded projects: grants, cooperative agreements, and contracts, such as the following:
            - Surveillance activities
            - Research activities
            - Transfer activities
            - Capacity-building activities

- Outputs (products of the research program–more details provided in Box 3-4):
    - Intramural:
        - Peer-reviewed publications, agency reports, alerts, and recommendations
        - Databases, websites, tools, and methods (including education and training materials)
        - Technologies developed and patents
        - Sponsored conferences and workshops
    - Extramural:
        - Program announcements
        - Requests for applications

- Intermediate outcomes:
    - Standards or guidelines issued by other agencies or organizations based in whole or in part on NIOSH research
    - Adoption and use of control or personal protective technologies developed by NIOSH
    - Evidence of industry, employer, or worker behavioral changes in response to research outputs
    - Use of NIOSH products by workers, industry, occupational health and safety professionals, healthcare providers, and others (including internationally)
    - NIOSH website hits and document requests
    - Unique staff or laboratory capabilities that serve as a national resource
    - Other intermediate outcomes, including those from extramural activities

- End outcomes:
    - Data on program impact on rates and numbers of injuries, illnesses, and hazardous exposures in the workplace (including trend data, if available)
    - Documentation of workplace risk reduction (quantitative, qualitative, or both)
- Description of current processes for setting research priorities and identifying emerging issues in the workplace

tp://www.nap.edu/catalog/12639.html

Many NIOSH programs have been evaluated by internal and other external bodies as part of an overall assessment of NIOSH, such as the Performance Assessment Rating Tool (PART) review,[3] or through evaluations of specific research program elements. The evaluation committees were asked to review, all prior evaluations of the program as an aid to understanding the evolution of the program and its elements. The National Academies committee evaluations, however, were independent of prior reviews and evaluations.

2. Assess External Factors

As depicted in the logic model (Figure 3-1), reductions in work-related injury and illness or in hazardous exposures (end outcomes) are dependent on stakeholder activities (external factors). Actions beyond NIOSH's control by those in labor, industry, regulatory entities, and elsewhere are necessary for NIOSH program activities to produce changes in end outcomes. Implementation of research findings may depend on existing or future policy considerations, economic conditions, and the public agenda.

External factors were considered as forces beyond the control of the NIOSH program that may affect the evolution of the program. External factors influence progress through all phases of the logic model, from inputs to end outcomes (see Figure 3-1). Identification of external factors by an evaluation committee is essential because it provides the context for evaluation of the program. External factors may be best assessed on the basis of the expert judgment of evaluation committee members who have knowledge of the field of research. NIOSH program staff provided their ideas on external factors early in the evaluation process. Information regarding external factors was also sought from other NIOSH program and management staff, OSHA and MSHA staff, and from other external stakeholders. Additionally, each evaluation committee chose other approaches to assess external factors. Factors external to a program might help or hinder the achievement of certain outcomes or might pre-

[3]PART focuses on assessing program-level performance and is one of the measures of success of the budget and performance integration initiative of the President's management agenda (*http://www.whitehouse.gov/omb/expectmore/summary/10002160.2004.html [accessed January 30, 2009]*).

tp://www.nap.edu/catalog/12639.html

sent formidable obstacles. The evaluation committees addressed both possibilities.

Some external factors may constrain research activities related to specific target populations, methodological issues, or resource availability. Evaluation committees examined whether

- Projects addressing a critical health need are technologically feasible. A workforce of appropriate size and with appropriate duration and distribution of exposure for measuring a health effect may not exist; for example, no population of workers has been exposed for 30 years to formaldehyde at the current OSHA permissible exposure limit (PEL), so the related cancer mortality cannot yet be directly assessed.
- Research is inhibited because NIOSH investigators are unable to access an adequate study population. Under current policy, NIOSH must either obtain an invitation by management to study a workplace or seek a judicial order to provide authority to enter a worksite (cooperation under court order may well be insufficient for effective research).
- Research is inhibited because the work environment, materials, and historical records cannot be accessed even with management and workforce cooperation.
- Adequate or established methods exist for assessing the environment.
- The NIOSH contribution to a particular field of research is reduced or difficult to estimate because other institutions are working in the same field.
- NIOSH resources are inadequate to tackle key questions.

Evaluation of the impact of NIOSH research outputs on worker health and safety also required consideration of external factors that might impede or aid implementation, measurement, and so on. Evaluation committees considered whether the following conditions exist and if so how this influences the research that NIOSH undertakes:

- Regulatory changes and implementation are unachievable because of obstacles to regulation or because of differing priorities of the regulatory agencies. For example, there may be no implementation of recommendations for improved respiratory protec-

tp://www.nap.edu/catalog/12639.html

tion programs for healthcare workers because of the lack or weakness of enforcement policies.

- A feasible control for a known risk factor or exposure has not been implemented because the cost of implementation is too high or because current economic incentives do not favor such actions.
- End outcomes are unobservable because baseline and continuing surveillance data are not available. For example, the current incidence of occupational noise-induced hearing loss is not known, although surveillance for a substantial threshold shift is feasible. (NIOSH conducts surveillance of some types of work-related illnesses, injuries, and hazards, but comprehensive surveillance is not possible with existing resources.)
- Reductions in adverse effects of chronic exposure cannot be measured. For example, 90 percent of identified work-related mortality is from diseases such as cancer that arise only after decades of latency after first exposure to a carcinogen. Effects of reducing exposure to a carcinogen therefore cannot be observed in the time frame of most interventions.
- A promulgated regulation requires a technology that has been developed but is not widely used.

3. Identify Time Frame to Be Evaluated

The NIOSH research program and other sources provided each evaluation committee with the history of the research program being evaluated and information on its major subprograms, goals, objectives, resources, and other pertinent information. Having that information allowed the committee to choose the time period most appropriate for the evaluation, with a focus on evaluating the program during the most recent appropriate period. For purposes of the eight reviews already completed, the evaluation committees considered three general timeframes: 1970–1995 (pre-NORA period), the period from the founding of NIOSH to the initiation of the National Occupational Research Agenda (NORA); 1996–2005 (NORA 1 period); and after 2005 (NORA 2 period). The period chosen for review took into consideration suggestions from the NIOSH research program under review. It was recognized that many of the intermediate and end outcomes documented in the selected time

tp://www.nap.edu/catalog/12639.html

frame are consequences of research outputs completed before that time period.

4. Identify Major Occupational Health and Safety Challenges in Program Area

Early in the assessment process, each evaluation committee identified, independently of NIOSH, the major occupational health and safety challenges for the research area being examined (Box A in Figure 3-2). In arriving at a list of challenges, the evaluation committees relied on surveillance findings, NIOSH investigations of sentinel events[4] (through health-hazard or fatality-assessment programs), external advisory inputs, and their own expert judgment. The evaluation committee then compared its own assessment of the challenges with the program's goals and objectives as outlined in the next step. The congruence between the two was useful during the assessment of relevance. In identifying and discussing the challenges, the evaluation committee included examples of best practices or described the components of the committee's vision of an ideal program.

5. Analyze Program Goals and Objectives

The research program's goals and objectives were evaluated with a focus on how each program goal is related to agency-wide strategic goals and to the program challenges (Box B in Figure 3-2). NIOSH research programs should be designed to be responsive to present or future workplace safety and health issues, and the evaluation committee was asked to provide an assessment of whether the program's goals and objectives are consistent with those issues. The evaluation committees recognized that NIOSH research priorities are sometimes circumstantial (e.g., congressionally mandated) rather than based on NIOSH's assessment of the state of knowledge.

[4]An occupational sentinel event is a disease, disability, or untimely death, which is occupationally related and whose occurrence may provide the impetus for further study or for the need to intervene (Rutstein et al., 1983).

tp://www.nap.edu/catalog/12639.html

Questions Considered in the Evaluation of Program Goals and Objectives:

1. Are the goals and objectives of the program well defined and clearly described?
2. How were the goals and objectives derived (or updated) through strategic planning processes?
3. How well aligned were program goals and objectives with NORA 1 priorities during the past decade?
4. How are current program goals and objectives related to current NIOSH goals?
5. Are the program's goals and objectives relevant to the major challenges for the research program and likely to address emerging issues in that specific research area (as determined by the evaluation committee)?
   - Did past program goals and objectives (as reflected in prior research and dissemination and transfer activities) focus on the most relevant problems and anticipate emerging issues?
   - Do the current program goals and objectives target the most relevant problems?

Assessment of Program Goals and Objectives

The evaluation committee was asked to provide a qualitative assessment discussing the relevance of the program's goals and objectives in relation to its major challenges.

6. Identify Major Program Components

Each evaluation committee determined how to disaggregate a program to achieve a manageable and meaningful evaluation of its components and of the overall program. Usually the disaggregation followed the strategic goals that the program identified. Although the research programs are built around intramural efforts, all relevant extramural efforts must be considered.

tp://www.nap.edu/catalog/12639.html

7. Evaluate Program Inputs, Activities, Outputs, and Outcomes

7a. Assess Inputs (Box C in Figure 3-2)

Planning inputs include input from stakeholders, surveillance and intervention data, and risk assessments. Production inputs include intramural and extramural funding, staffing, management structure, and physical facilities.

The evaluation committee examined existing intramural and extramural resources and, potentially, prior surveys or case studies that might have been developed specifically to assess progress in reducing workplace illnesses and injuries and to provide information relevant to the targeting of research to future needs. The NIOSH research program provided the evaluation committee with relevant planning and production inputs (see below and Box 3-2 for examples).

Planning inputs. Planning inputs can be qualitative or quantitative. Sources of qualitative inputs for NIOSH included the following:

- Other NIOSH programs;
- Federal advisory committees, such as the Board of Scientific Counselors, the Mine Safety and Health Research Advisory Committee, and the National Advisory Committee on Occupational Safety and Health;
- NORA research partners and stakeholders, NORA strategic research plans, and the NORA Liaison Committee and federal liaison committee recommendations;
- Industry, labor, academe, professional associations, industry associations, and the Council of State and Territorial Epidemiologists; and
- OSHA and MSHA strategic plans and other federal research agendas.

Attention was given to how comprehensive the inputs have been and to what extent gaps in input have been identified and considered by the program being evaluated. Sources of quantitative inputs for NIOSH included the following:

tp://www.nap.edu/catalog/12639.html

- Intramural surveillance information, such as descriptive data on exposures and outcomes (appropriate data may be available from a number of NIOSH divisions and laboratories);
- Reports from the NIOSH Health Hazard Evaluation (HHE) program;
- Reports from the Fatality Assessment and Control Evaluation (FACE) program;
- Extramural health-outcome and exposure-assessment data from OSHA, MSHA (both safety and health inspection data), the Bureau of Labor Statistics, the U.S. Department of Defense (DoD), and the U.S. Department of Agriculture (USDA) (fatality, injury, and illness surveillance data); state government partners, including NIOSH-funded state surveillance programs, such as the Sentinel Event Notification System for Occupational Risks, Adult Blood Lead Epidemiology and Surveillance, and state-based FACE programs; and nongovernmental organizations, such as the National Safety Council, the Association of Occupational and Environmental Clinics (AOEC), the American Society of Safety Engineers, and the American College of Occupational and Environmental Medicine; and
- Appropriate data from investigator-initiated extramural research funded by NIOSH.

Production inputs. For the research program under review, NIOSH program staff identified portions of the NIOSH intramural budget, staff, facilities, and management that played major roles in the research program. Production inputs were described primarily in terms of support for intramural research projects, relevant extramural projects (particularly cooperative agreements and contracts), HHEs that supported program goals, and related staffing levels. Consideration was also given to leveraged funds provided by partners such as the National Institutes of Health (NIH) and the Environmental Protection Agency (EPA) for joint requests for applications or program announcements and to OSHA, MSHA, and DoD contracts as well as collaborations with other NIOSH programs.

Using this evaluation model, assessment of inputs included the evaluation committee's consideration of the degree to which allocation of funding and personnel was commensurate with the resources needed to conduct the research and the extent to which funding for the relevant intramural research activity has been limited by lack of discretionary

tp://www.nap.edu/catalog/12639.html

spending beyond salaries, such as travel, supplies, and external laboratory services.

Questions considered in the evaluation of inputs:

1. Are planning and production inputs consistent with program goals?
2. How well are major planning and production inputs used to support the major activities?
3. Is input obtained from external stakeholders, including from those representing vulnerable working populations and small businesses?
4. Are production inputs (intramural and extramural funding, staffing, management, and physical infrastructure resources) consistent with program goals and objectives?

Assessment of inputs. The evaluation committee was asked to provide a qualitative assessment that discussed the quality, adequacy, and use of inputs.

7b. Assess Activities (Box D in Figure 3-2)

Activities are defined as the efforts of program staff, grantees, and contractors. Activities of the NIOSH program under review were divided into research and transfer activities. Box 3-3 suggests the types and organization of information that can be useful in evaluating program activities. Some types of research activity may not be applicable to a given NIOSH program. Research activities include surveillance, health-effects research, exposure assessment research, safety-design and safety-systems research, intervention research, and diffusion and dissemination research. Transfer activities include marketing analysis, information dissemination, training, technical assistance, and technology transfer. Depending on the scope of the program under review, activities may also be grouped by research program objectives or by subprograms.

Conventional occupational safety and health research efforts appropriately focus on injury, illness, or death; on biomarkers of exposure; and on health effects of new technology, personal protective equipment, and regulations. Consideration was also given to the types of surveillance data needed. Assessment of the program's activities relevant to socio-

tp://www.nap.edu/catalog/12639.html

economic and policy research and diffusion research were also considered because these research endeavors can provide information needed to effect important outcomes farther out on the causal chain to influence health and safety in the workplace. Examples of other types of research that could have been useful to the evaluation committees in examining activities relevant to the program's mission included the following:

- Surveillance research to assess the degree of significant or systematic underreporting of relevant injuries, illnesses, and biomarkers;
- Socioeconomic research on cost shifting between workers' compensation and private insurance;
- Research on methods to build the health and safety capacity of primary care clinicians in community health centers and other healthcare settings to improve the recognition and treatment of work-related conditions;
- Transfer research on how to change the health and safety knowledge and behavior of adolescents to improve the likelihood of reduced injuries as they enter the workforce; and
- Community-based participatory research to explore how recent immigrants and workers employed for a longer time in the United States understand acceptable health and safety risks, with the purpose of better targeting the workforce training needs of immigrant workers.

BOX 3-3
Examples of NIOSH Program Research and Transfer Activities

Surveillance (including surveillance of injuries, illnesses, and hazards)

Health-effects research (illnesses, injuries, and biomarkers):
- Epidemiology
- Toxicology
- Physical and safety risk factors (laboratory based)
- Development of clinical screening methods and tools

Exposure assessment research:
- Chemical hazards
- Physical hazards
- Biologic hazards
- Ergonomic hazards
- Safety (traumatic injury) hazards

tp://www.nap.edu/catalog/12639.html

Safety design and safety systems research

Intervention research:
- Control technologies
- Engineering controls and alternatives
- Administrative controls
- Personal protective equipment
- Work organization
- Community participation
- Policy (e.g., alternative approaches to targeting inspections)
- Design for safety
- Emergency preparedness and disaster response

Diffusion and dissemination research:
- Training effectiveness
- Information dissemination effectiveness
- Diffusion of technology

Health services and other research:
- Access to occupational health care
- Infrastructure–delivery of occupational health services
- Socioeconomic consequences of work-related injuries and illnesses
- Workers' compensation

Technology transfer and other transfer activities:
- Information dissemination
- Training programs
- Marketing analysis
- Technical assistance

Transfer activities were assessed to determine whether the program appropriately targets its outputs in a manner that will have the greatest impact. Ideally, information dissemination should be proactive, and strategic dissemination should be informed by research on the diffusion of new technologies, processes, and practices. Highly relevant information and technology transfer include plans for transfer to all appropriate worker populations, including those considered vulnerable. Training should be incorporated into the strategic goals of all research fields where appropriate.

The evaluation committee reviewed project-level research and transfer activities (including surveillance activities) that have been completed, are in progress, or are planned by the program under review. Programs were asked to provide a list of activities and specify whether the activities were intramural or extramural.

tp://www.nap.edu/catalog/12639.html

The evaluation committee assessed each research activity outlined in Box 3-3 that is or should be an important element of the specific program being evaluated. In the case of a sector-based research program (e.g., mining or construction) for which health-effects research is not being evaluated, each committee determined what research activities were consistent with the program's goals and objectives, and then assessed the value of the activities.

Questions considered in assessing research activities:

1. What are the major subprograms or groupings of activities within the program?
2. Are activities consistent with program goals and objectives?
3. Are research activities relevant to the major challenges of the research program?
   - Do they address the most serious outcomes?
   - Do they address the most common outcomes?
   - Do they address the needs of both genders, vulnerable working populations, and small businesses?
4. Are NIOSH research activities pioneering in opening new and useful fields of research to be further explored by NIOSH and others?
5. Are research activities appropriately responsive to the input of stakeholders?
6. To what extent do research activities involve external partnerships?
7. Are partners involved early in the research process to allow them to participate in determining research objectives and research design?
8. Were original resource allocations appropriate for the research activities, and do they remain appropriate?
9. To what extent do peer reviews (internal, external, and midcourse) affect the activities?
10. Is there adequate monitoring of quality assurance procedures to ensure credible research data, analyses, and conclusions?

tp://www.nap.edu/catalog/12639.html

Questions considered in assessing transfer activities:

1. Is a coherent program of transfer activities planned?
2. Have staff thought through issues of compatibility, cost, and simplicity in designing information and transfer products?
3. Are the program's publications and information dissemination, training, education, and technical assistance efforts successful in reaching the workplace or relevant stakeholders in other settings? How widespread is the response?
4. To what degree have stakeholders responded to program information and training products?
5. Is there evidence that the formats for information products were selected in response to stakeholder preferences?
6. To what extent do program personnel rely on assessment of stakeholder needs and reactions to prototype information and training projects (formative evaluation techniques)?
7. To what extent does the program build research and education capacity internally and among stakeholders?

Assessment of activities. Each evaluation committee was asked to provide a qualitative assessment of the relevance of these efforts. This assessment included consideration of the external factors that constrained choices of research projects and the relevance and effectiveness of transfer activities. The evaluation committee considered the appropriateness of resource allocations. A highly relevant program would address high-priority needs, produce high-quality results, be appropriately collaborative, be of value to stakeholders, and be substantially engaged in transfer activities. A program might be less relevant to the extent that those key elements were not up to the mark or were missing. The committee's discussions covered those aspects in sufficient detail to arrive at a qualitative assessment of the activities. Assessment of the transfer activities included considerations of program planning, coherence, and impact. The evaluation committee also considered the incorporation of international research results into knowledge-transfer activities conducted by the NIOSH program for U.S. industry sectors.

tp://www.nap.edu/catalog/12639.html

7c. Assess Outputs (Box E in Figure 3-2)

For the NIOSH evaluations, an output is a direct product of a NIOSH research program. Outputs may be designed for researchers, practitioners, intermediaries, and end-users, such as employers and employees. Outputs can include publications in peer-reviewed journals, recommendations, reports, website content, workshops and presentations, databases, educational materials, scales and methods, new technologies, patents, and technical assistance. Outputs of the research program's extramurally funded activities were also considered. Box 3-4 lists examples of major outputs considered by the evaluation committees. Each NIOSH research program was asked to make every effort to include all pertinent data of the types listed in Box 3-4 in the materials submitted to the committee.

Outputs may be tailored to the intended audience to communicate information most effectively and increase the likelihood of comprehension, knowledge, attitude formation, and behavioral intent. The extent of use of formative evaluation data and the extent of user feedback in the design of the output can be considered indicators of appropriate quality assessment.

BOX 3-4
Examples of Research Program Outputs

Peer-reviewed publications by NIOSH staff:

- Number of original research articles by NIOSH staff and citations
- Number of review articles by NIOSH staff (including best-practices articles) and citations
- Publications in the field of interest with other support by investigators also funded by NIOSH (e.g., ergonomic studies with other support by an investigator funded by NIOSH to do ergonomics work, in which case NIOSH should get some credit for seeding interest or drawing people into the field)

Peer-reviewed publications by external researchers funded by NIOSH:

- Number of NIOSH-funded original research articles by external researchers and citations
- Number of NIOSH-funded review articles by external researchers (including best-practices articles) and citations
- Collaboration with other government or academic researchers

NIOSH reports:

- Number of written reports and citations

tp://www.nap.edu/catalog/12639.html

Sponsored conferences and workshops:

- Number of sponsored conferences
- Number of sponsored workshops
- Description of conferences and workshops (title, date, sponsors, target audience, number of participants, and resulting products)

Databases:

- Number of major databases created by NIOSH staff
- Number of major databases created by external researchers funded by NIOSH grants
- Description of databases:
  - Title, objective (in one to four sentences), and start and stop dates
  - Partial versus complete sponsorship (if partial, who were co-sponsors?)
  - Study or surveillance system design, study population, and sample size
  - Primary "products" of the database (e.g., number of peer-reviewed articles and reports)

Recommendations:

- Number of major recommendations
- Description of recommendations:
  - Complete citation (article, report, or conference where recommendation was made)
  - Summary in one to four sentences
  - Percentages of target audiences and decision makers that have adopted the recommendation (up to 10 years after release)
  - Examples of implementation in the field

Tools, methods, and technologies:

- Number of major tools, methods, and technologies (includes training and education materials)
- Descriptions
  - Title and objective (in one to four sentences)
  - Complete citation (if applicable)
  - Percentage of target audience that has used the tools, methods, or technologies (up to 10 years after release)
  - Up to three examples of implementation in the field

Patents:

- Total number of patents
- For each:
  - Title and objective (in one to four sentences)
  - Complete citation
  - Percentage of target audience that has used product (up to 10 years after release)
  - Up to three examples of implementation in the field

tp://www.nap.edu/catalog/12639.html

Miscellaneous:
- Any other important program outputs

Activities such as collaborations can also legitimately be conceptualized as outputs, because the collaboration itself is a result of NIOSH efforts. Cooperation, coordination, more intensive collaboration, and eventual formal partnering can be considered important outputs leading to desirable intermediate outcomes. Technology transfer and knowledge transfer are facilitated significantly through such relationships. The extent of collaboration with other organizations in the determination of research agendas, the conduct of research, the dissemination of research results, and interorganizational involvement in the production of outputs may be measures of output quality and quantity. The evaluation committees considered coauthorship while determining the importance of research by the NIOSH program to the broader research community.

The NIOSH program was asked to provide information on all relevant outputs of the program under review that were produced during the chosen period.

Questions considered in the evaluation of outputs:

1. What are the major outputs of the research program?
2. Are output levels consistent with resources allocated (were resources allocated and used efficiently to produce outputs)?
3. Does the research program produce outputs that address high-priority areas?
4. To what extent does the program generate important new knowledge or technologies?
5. Do any widely cited, peer-reviewed publications report "breakthrough" results?
6. What, if any, internal or external capacity-building outputs are documented?
7. Are outputs relevant to both genders and vulnerable populations, and do they address the needs of small businesses?
8. Are products user-friendly with respect to readability, simplicity, and design?
9. To what extent does the program help to build the internal or extramural institutional knowledge base?

tp://www.nap.edu/catalog/12639.html

10. Does the research produce effective cross-agency, cross-institute, or internal–external collaborations?
11. To what extent does the program build research and education capacity (internal or external)?

Assessment of outputs. The evaluation committees were asked to provide a qualitative assessment, including discussion of relevance and utility. The outputs of a highly ranked program address needs in high-priority areas, contain new knowledge or technology that is effectively communicated, contribute to capacity building inside and outside the program, and are relevant to pertinent populations. The committees were asked to provide a discussion that covered those aspects in sufficient detail to support the qualitative assessment of the outputs.

7d. Assess Outcomes (Boxes F and G in Figure 3-2)

Intermediate outcomes. Intermediate outcomes are external stakeholder actions to which the program contributed. They reflect the impact of program activities and may lead to the desired end outcome of improved worker safety and health. Intermediate outcomes in the NIOSH evaluations included the production of guidelines or regulations based wholly or partly on NIOSH research by those outside of NIOSH (products adopted as public policy or as policy or guidelines by private organizations or industry); contributions to training and education programs sponsored by other organizations; use of publications or other materials by workers, industry, and occupational safety and health professionals in the field; secondary dissemination of program activities and outputs through trade and mass media coverage; and citations of NIOSH research by industrial and academic scientists.

Intermediate outcomes allow inference that a program's outputs are associated with observed changes in the workplace. Thus, an intermediate outcome reflects an assessment of worth by NIOSH program stakeholders (e.g., managers in industrial firms) about NIOSH research or its products (e.g., NIOSH training workshops). Intermediate outcomes that are difficult to monitor, but may be valid indicators of relevance or utility, include self-report measures by users of NIOSH outputs. Self-reported indicators include the extent to which key intermediaries find value in NIOSH products or databases for the repackaging of health and

tp://www.nap.edu/catalog/12639.html

safety information, the extent to which NIOSH recommendations are in place and attended to in workplaces, and employee or employer knowledge of and adherence to NIOSH-recommended practices.

Questions considered in the evaluation of intermediate outcomes:

1. Do program outputs result in or contribute to stakeholder training or education activities used in the workplace or in school or apprentice programs? If so, how?
2. Do program activities and outputs result in regulations, public policy, or voluntary standards or guidelines that are transferred to or created by the workplace?
3. Have the program's activities and outputs resulted in changes in employer or worker practices associated with the reduction of risk—for example, in the adoption of new feasible control or personal protective technologies or administrative control concepts?
4. Does the program contribute to changes in healthcare practices intended to improve recognition and management of occupational health conditions?
5. Do program activities and outputs result in research partnerships with stakeholders that lead to changes in the workplace?
6. To what extent do stakeholders find value in the NIOSH program's products, as shown by document requests, website hits, conference attendance, and similar evidence of stakeholder interest?
7. Does the program or a subprogram provide unique staff or laboratory capability that is a necessary national resource? If so, is it adequate, or does it need to be enhanced or reduced?
8. Have program activities and outputs resulted in interventions that protect both genders and vulnerable workers or address the needs of small businesses?
9. To what extent did the program contribute to increased capacity at worksites to identify or respond to safety and health threats?

Assessment of intermediate outcomes. The evaluation committees were asked to provide a qualitative assessment of product development, usefulness, and impact with consideration given to the relative value of

tp://www.nap.edu/catalog/12639.html

intermediate outcomes (the framework committee recommended applying the well-accepted hierarchy-of-controls model). Discussions could include comments on how widely products have been used or programs implemented. The qualitative discussion should be specific about the various products developed by the program and the extent of their use by certain entities (e.g., industry, labor, government) for specific purposes. Discussions included whether the products have resulted in changes in the workplace or in the reduction of risk; the recognition accorded to the program or the facilities by its peers (e.g., recognition as a "center of excellence" by national and international communities) was also considered in the assessment. To be highly ranked, a program should have high performance in most of the relevant questions in this section. One aspect of the evaluation was considering whether the same changes in stakeholder activities and behaviors probably would have occurred without efforts by the NIOSH program.

End outcomes. Each evaluation committee was asked to assess to the greatest extent possible the program's contribution to end outcomes—improvements in workplace health and safety (impact). For purposes of this evaluation, end outcomes are health- and safety-related changes that are a result of program activities, specifically, decreases in injuries, illnesses, and hazardous exposures or risks. Data on reductions in work-related injuries, illnesses, and hazardous exposures were available for a few of the programs, and in some cases they were quantifiable. When there was no direct evidence of improvements in health and safety, intermediate outcomes were used as proxies for end outcomes in assessing impact, and the evaluation committee qualified their findings. The evaluation committees described the realized or potential benefits of the NIOSH program.

Assessing the causal relationship between NIOSH research and specific occupational health and safety outcomes was seen as a major challenge because NIOSH does not have direct responsibility or authority for implementing its research findings in the workplace. Furthermore, the benefits of NIOSH research program outputs can be realized, potential, or limited to the knowledge gained. Studies that conclude with negative results may nevertheless have incorporated excellent science, close off unproductive areas of research, and contribute to the knowledge base. The generation of important knowledge is a recognized form of outcome in the absence of measurable impacts.

tp://www.nap.edu/catalog/12639.html

The impact of research, particularly applied research as conducted by NIOSH, depends on the existence of a "receptor" for research results, such as a regulatory agency, a professional organization, an employer, an employee organization, or in some cases, employees themselves. The evaluation committee was asked to consider issues related to the various stages that lead to outcomes, including the following:

- Did research by the program identify a gap in protection or a means of reducing risk?
- Did the program convey this information to potential users in a usable form?
- Were research results (e.g., recommendations, technologies) applied?
- Did the applied results lead to desired outcomes?

Quantitative data were preferable to qualitative, but qualitative analysis was necessary in some cases. Sources of quantitative data relevant to NIOSH included the following:

- Bureau of Labor Statistics (BLS) data on fatal occupational injuries (the Census of Fatal Occupational Injuries) and nonfatal occupational injuries and illnesses (the annual Survey of Occupational Injuries and Illnesses);
- NIOSH intramural surveillance systems, such as the National Electronic Injury Surveillance System, the coal worker X-ray surveillance program, and agricultural worker surveys conducted by NIOSH in collaboration with the USDA;
- State-based surveillance systems, such as the NIOSH-funded Adult Blood Lead Epidemiology and Surveillance Program and the Sentinel Event Notification System for Occupational Risks programs for asthma, pesticides, silicosis, noise-induced hearing loss, dermatitis, and burns; state-level vital statistics systems and other health data systems such as cancer registries and hospital discharge and emergency department datasets;
- Selected state worker compensation programs; and
- Exposure data collected in the OSHA Integrated Management Information System.

tp://www.nap.edu/catalog/12639.html

The framework committee was unaware of mechanisms for the surveillance of many occupation-related chronic illnesses, such as cancers that arise from long exposure to chemicals and other stressors. The incidence and prevalence of many such outcomes may best be evaluated by investigator-initiated research. Research that leads to new effective surveillance concepts or programs warrants special recognition.

The evaluation committees were asked to consider the strengths and weaknesses of outcome data sources. Quantitative accident, injury, illness, and employment data and databases are subject to error and bias and should be used by the evaluation committees only for drawing inferences after critical evaluation and examination of available corroborating data. For example, occupational illnesses are widely recognized as being poorly documented in the BLS Survey of Occupational Injuries and Illnesses, which captures only incident cases among active workers. Health practitioners often have difficulty in diagnosing the component of illnesses that may be related to work; furthermore, few practitioners are adequately trained to make such an assessment. Many of these illnesses have long latencies and do not appear until years after individuals have left the employment in question. Additionally, surveillance programs may systematically undercount some categories of workers, such as contingent workers.

In addition to measures of illness and injury, measures of exposure to chemical and physical agents and to safety and ergonomic hazards can be useful. Measures of exposure or probability of exposure can serve as an appropriate proxy for disease or injury when a well-described association exists between occupational exposure and health. In such instances, a decrease in exposure can be accepted as evidence that the end outcome of reduced illness or injury is being achieved. Such assumptions are particularly necessary when the latent period between exposure and disease outcome makes effective evaluation of the relevant end outcome infeasible, as in the case of asbestos exposure and lung cancer.

The reduction in the number of worksites that exceed an OSHA PEL or an American Conference of Governmental Industrial Hygienists threshold limit value is a quantitative measure of improvement of occupational health awareness and reduction of risk. In addition to exposure level, the number of people exposed and the distribution of exposure levels can provide quantitative data. Other evidence includes air monitoring data, reduction in requirements for use of personal protective equipment, and reduction in ergonomic risks.

tp://www.nap.edu/catalog/12639.html

Challenges posed by inadequate or inaccurate measurement systems should not drive programs out of difficult fields of study, and the evaluation committees should be aware of such possibilities. In particular, contingent and informal working arrangements that place workers at greatest risk are also those on which surveillance information is almost totally lacking, so novel methods for measuring impact may be required.

The commitments of industry, labor, and government to health and safety are critical external factors. Several measures of this commitment can be useful for the evaluation committee: monetary commitments, attitude, staffing, and surveys of relative importance. To the extent that resources allocated to safety and health are limiting factors, the evaluation committee also explicitly assessed the performance of the NIOSH program in the context of constraints.

Questions considered in the evaluation of end outcomes:

1. What are the amounts and qualities of relevant end outcome data, including data documenting injuries, illness, exposure, and productivity affected by health?
2. What are the temporal trends in the data?
3. Is there objective evidence of improvement in occupational safety or health?
4. To what degree is the NIOSH program or subprogram responsible for improvement in occupational safety or health?
5. How do findings compare with data from comparable groups in the United States or the corresponding populations in other countries?
6. What is the evidence that external factors have affected outcomes or outcome measures?
7. Has the program been responsible for changes in outcomes outside the United States?

Assessment of end outcomes. The evaluation committee was asked to provide a qualitative assessment of program impact, discussing the evidence of reductions in injuries and illnesses or their appropriate proxies.

tp://www.nap.edu/catalog/12639.html

Other outcomes. Regarding the NIOSH study, there may be as-yet unappreciated health and safety impacts or other beneficial social, economic, and environmental outcomes as a result of NIOSH activities. NIOSH study results may be influential outside the United States, and there may be evidence of implementation of NIOSH recommendations and training programs abroad.

Questions considered in the evaluation of other outcomes:

1. Is the program likely to produce a favorable change that has not yet occurred or not been appreciated?
2. Has the program been responsible for social, economic, security, or environmental outcomes?
3. Have program activities and outputs impacted occupational health and safety in other countries?

Assessment of other outcomes. The evaluation committees were asked to consider other outcomes, including beneficial changes that are expected to occur; social, economic, security, or environmental outcomes; and the impact that the program has had on international occupational safety and health.

8. Determine Scores for Relevance and Impact and Provide the Rationale

The evaluation committees assigned an integer score for the *relevance* of the research program to the improvement of occupational safety and health and another integer score for the *impact* of the program on such improvements. Using their expert judgment, the committees rated the relevance and impact of the overall research program by first summarizing their assessments of the major goals or subprograms and then appropriately weighting the goal areas (or subprograms) to determine the overall program ratings.[5]

[5]In light of substantial differences among the types of research programs, the framework committee chose not to construct a single algorithm to use in weighting the goals or subprograms.

tp://www.nap.edu/catalog/12639.html

Relevance and impact scores were based on 5-point categorical scales established by the framework committee (described below), in which 1 is the lowest and 5 the highest rating. The framework committee made an effort to establish mutually exclusive rating categories in the scales. The evaluation committee determined how individual goal areas (or subprograms) influenced final scores. Final program ratings consisted of integer scores for relevance and impact and prose justification of the scores.

Box 3-5 provides an overview of the issues to be considered in determining ratings of relevance and impact. Evaluation committees were asked to consider items 1 through 4 in Box 3-5 for the overall program and to assess the relevance of the program by reviewing the committee's responses to the questions evaluating the program's challenges, goals and objectives, inputs, activities, and outputs (Section 7). The evaluation committee evaluated separately the extent to which the program's research efforts are in high-priority areas and the extent to which the program is involved in transfer activities. Transfer activities occur in two contexts: (1) efforts by the NIOSH program to translate intellectual products into practice and (2) stakeholder efforts to integrate NIOSH results into the workplace.

To assess impact, each evaluation committee first needed to consider the available evidence of changes in work-related risks and the adverse effects and external factors related to the changes. The evaluation committee reviewed the responses to the questions on the reviews of outputs, intermediate outcomes, and end outcomes and systematically assessed the impact of the research program. Items 2 to 7 in Box 3-5 address these areas. The evaluation committee needed to judge, for example, whether outcomes occurred earlier than they would have or are better than they would have been in the absence of the research program, or whether outcomes would have occurred were it not for external factors beyond the control of the NIOSH program.

Scoring of Relevance

As discussed in previous sections, numerous factors may be considered in assessing relevance. The scoring criteria focus on the evaluation committee's assessment of whether the program appropriately set priorities among research needs as well as how engaged the program was in

tp://www.nap.edu/catalog/12639.html

BOX 3-5
Overview of the Issues

Assess the following for each program:

1. Relevance of current and recently completed research and transfer activities to objective improvements in workplace safety and health.
2. Contributions of the NIOSH program's research and transfer activities to changes in work-related practices and reduction in workplace exposures, illnesses, or injuries.
3. Contributions of the NIOSH program's research and transfer activities to improvements in work-related practices.
4. Contributions of the NIOSH program's research to productivity, security, or environmental quality (beneficial side effects).
5. Evidence of policy, technological, behavioral, and other changes that would reduce risk in the workplace (intermediate outcomes).
6. Evidence of reduction in workplace exposures, illnesses, or injuries (end outcomes).
7. Evidence of external factors that prevented translation of NIOSH research results into intermediate or end outcomes.

appropriate transfer activities to move research findings into the workplace. Since the evaluation of NIOSH programs included assessment of research activities and knowledge transfer activities, both are considered in the final relevance score. With respect to research, the key indicator is the extent to which the program's research is in priority subject areas (high priority, priority, lesser priority, or not focused on priorities); with respect to transfer, the key indicator is the level of engagement in appropriate transfer activities (in this case, significantly engaged, engaged, or not engaged). This approach resulted in a complex scoring system that tries to address the best and worst cases and any variations in between. Box 3-6 lists the criteria for scoring the overall relevance of the NIOSH research program. Table 3-1 provides guidance regarding how the committee may weight research priorities and transfer levels when determining relevance scores.

The evaluation committee considered both completed research and research that is in progress in its assessment of relevance. The committee kept in mind how well the program has considered the frequency and severity of the problems being addressed; whether appropriate attention has been directed to issues regarding both sexes, vulnerable populations, or hard-to-reach workplaces; and whether the different needs of large and small businesses have been considered. Each committee determined how to consider external factors in assigning program scores.

tp://www.nap.edu/catalog/12639.html

BOX 3-6
Scoring Criteria for Relevance

5 = Research is in high-priority subject areas and the NIOSH program is significantly engaged in appropriate transfer activities for completed research projects/reported research results.
4 = Research is in high-priority subject areas and the NIOSH program is engaged in appropriate transfer activities for completed research projects/reported research results; or research is in priority subject areas and the NIOSH program is significantly engaged in appropriate transfer activities for completed research projects/reported research results.
3 = Research is in high-priority subject areas, but the NIOSH program is not engaged in appropriate transfer activities; or research is in priority subject areas but the NIOSH program is not significantly engaged in appropriate transfer activities; or research focuses on lesser priorities but the NIOSH program is significantly engaged in appropriate transfer activities.
2 = Research program is focused on lesser priorities and the NIOSH program is not significantly engaged in appropriate transfer activities.
1 = Research program is not focused on priorities.

TABLE 3-1 Guidance for Weighting Research Priority and Transfer Activities

Assessment of Research Priority	Engagement in Appropriate Transfer Activities	Applicable Score
High priority	Significantly engaged	5
High priority	Engaged	4
Priority	Significantly engaged	4
High priority	Not engaged	3
Priority	Engaged or not engaged	3
Lesser priority	Significantly engaged	3
Lesser priority	Engaged or not engaged	2
Not focused on priorities	Any level of engagement	1

tp://www.nap.edu/catalog/12639.html

Scoring of Impact

Box 3-7 provides the criteria established for the rating of impact. The evaluation committee primarily considered completed research outputs. In assigning a score for impact, it is important to recognize that a "major contribution" (required for a score of 5) does not imply that the NIOSH program was solely responsible for observed improvements in worker health and safety. Many factors may be required to effect improvements. The committee could say that the NIOSH program made "major contributions" if the improvements would not have occurred when they did without the program's efforts.

The framework committee had some concern that the imposed scoring criteria for impact might be considered a promotion of the conventional occupational-health research paradigm that focuses on health-effects and technology research without much emphasis on the socioeconomic, policy, surveillance, and diffusion research (as opposed to diffusion activities) needed to effect change. The evaluation committees were asked to remember that not all intermediate outcomes occur in the workplace. Important outcomes that NIOSH can affect also occur much farther out on the causal chain. NIOSH, for example, has an important role in generating knowledge that may contribute to changing norms in the insurance industry, in health-care practice, in public health practice, and in the community at large.

BOX 3-7
Scoring Criteria for Impact

5 = Research program has made major contribution(s) to worker health and safety on the basis of end outcomes or well-accepted intermediate outcomes.

4 = Research program has made some contributions to end outcomes or well-accepted intermediate outcomes.

3 = Research program activities are ongoing and outputs are produced that are likely to result in improvements in worker health and safety. Well-accepted outcomes have not been recorded.

2 = Research program activities are ongoing and outputs are produced that may result in new knowledge or technology, but only limited application is expected. Well-accepted outcomes have not been recorded.

1 = Research activities and outputs do not result in or are not likely to have any application.

NA = Impact cannot be assessed; program is not mature enough.

tp://www.nap.edu/catalog/12639.html

The evaluation committees considered whether some of those issues need to be addressed and considered as external factors that facilitate or limit application of more traditional research findings. Given the rapidly changing nature of work and the workforce and the intractable problems in manufacturing, mining, and other fields, the evaluation committees were encouraged to think beyond the conventional paradigm.

9. Assess the NIOSH Process for Targeting Priority Research Needs and Provide the Committee's Assessment of Emerging Issues

Among the most challenging aspects of research in illness and injury prevention are the identification of new or emerging needs and trends and the formulation of a research response that uses scarce resources to best effect in anticipation of them.

The second charge to the evaluation committee was assessment of the research program's effectiveness in targeting new research and identifying emerging issues in occupational safety and health most relevant to future improvements in workplace protection. The evaluation committee was asked to provide a qualitative narrative assessment of the program's process for determining priorities for research and emerging workplace issues. The committee also independently identified emerging workplace issues that the NIOSH program should be prepared to address.

The evaluation committees reviewed the procedures that the NIOSH program has in place to identify needed research relevant to the NIOSH mission and reviewed the success that the NIOSH program has had in identifying and addressing research related to emerging issues. For example, the program should be involved in examining leading indicators from other federal agencies (e.g., the EPA, the Department of Labor, the National Institute of Standards and Technology, the NIH, the DoD, and the Department of Commerce) that track or provide data on new technologies, new products, new processes, and disease or injury trends.

The NIOSH HHE program offers a potential source for the identification of emerging research needs. The evaluation committee needed to determine whether the program under review appropriately considered pertinent HHE investigation findings. Additional emerging issues may have been revealed through consideration of NIOSH and NIOSH-funded FACE reports, AOEC reports, U.S. Chemical Safety Board investigations, and the Sentinel Event Notification System for Occupational Risks

tp://www.nap.edu/catalog/12639.html

and other state-based surveillance programs. Appropriate federal advisory committees and other stakeholder groups were also consulted to provide qualitative information.

The evaluation committee systematically assessed how the research program targets new research by evaluating each goal area or subprogram for the items listed in Box 3-8.

Questions Considered in Identifying Emerging Issues:

1. What information does the NIOSH program review to identify emerging research needs?
   - What is the process for review?
   - How often does the process take place?
   - How are NIOSH staff scientists and leadership engaged?
   - What is the process for moving from ideas to formal planning and resource allocation?
2. How are external stakeholders involved?
   - What advisory or stakeholder groups are asked to identify emerging research targets?
   - How often are such groups consulted, and how are suggestions followed up?
3. What new research targets have been identified for future development in the program under evaluation?
   - How were they identified?
   - Were lessons learned that could help to identify other emerging issues?
   - Does the evaluation committee agree with the issues identified and selected as important and with the NIOSH program's response, or were important issues overlooked?
   - Is there evidence of unwise expenditure of resources on unimportant issues?

The evaluation committee members used their expert judgment both to evaluate the emerging research targets identified by the NIOSH program and to provide recommendations on improvements to the program or additional research that NIOSH had not yet identified. Recommendations included a brief statement of their rationale.

tp://www.nap.edu/catalog/12639.html

BOX 3-8
Targeting of New Research and Identification of Emerging Issues

Assess the following:

1. Past and present effectiveness in targeting most relevant research needs.
2. Effectiveness in targeting research in fields most relevant to future improvements in occupational safety and health.

10. Prepare Report by Using the Template Provided as a Guide

Consistency and comparability among evaluation committee report formats was desirable, but the framework committee recognized that each NIOSH research program is different and that each evaluation committee was independent. The outline provided in Box 3-9 flows from the framework committee's review of the generalized NIOSH logic model (Figure 3-1) and the overviews of the evaluation process (Box 3-1, Figure 3-2). The evaluation committees were free to use or adapt this outline as necessary when organizing their final reports. The framework committee encouraged each evaluation committee to look at prior evaluation committee reports for organizational ideas.

BOX 3-9
Suggested Outline for Evaluation Committee Reports

I. Introduction

This section should be a brief descriptive summary of the history of the program being evaluated with respect to pre-NORA, NORA 1, and current and future plans of the research program presented by the NIOSH program. It should present the context for the research on safety and health; goals, objectives, and resources; groupings of goal areas or subprograms; and any other important pertinent information (a list of the NIOSH materials reviewed should be provided in Appendix C.)

II. Evaluation of the Program (Charge 1)

A. Evaluation summary: should include a brief summary of the evaluation with respect to impact and relevance, scores for impact and relevance, and summary statements
B. Strategic goals and objectives: should describe assessment of the extent to which program strategic plans reflect program relevance

tp://www.nap.edu/catalog/12639.html

C. Assessment of inputs: should describe adequacy of inputs to achieve goals
D. Assessment of activities: should describe assessment of the relevance of the activities
E. Assessment of research program outputs: should describe assessment of relevance and potential usefulness of the research program's outputs
F. Assessment of intermediate outcomes and causal impact: should describe assessment of the intermediate outcomes and the program's contribution to them; should include the likely impacts and recent outcomes in the assessment.
G. Assessment of end outcomes: should describe the end outcomes related to health and safety and provide an assessment of the type and degree of attribution to the NIOSH program
H. Assessment of other outcomes: should discuss health and safety impacts that are expected to occur; beneficial social, economic, and environmental outcomes; and international dimensions and outcomes
I. Summary of ratings and rationale

III. NIOSH Targeting of New Research and Identification of Emerging Issues (Charge 2)

The evaluation committee should assess the progress that the NIOSH program has made in targeting new research in occupational safety and health. The evaluation committee should assess whether the NIOSH program has identified important emerging issues that appear especially important in terms of relevance to the mission of NIOSH. The evaluation committee should respond to NIOSH's perspective and add its own recommendations.

IV. Recommendations for Program Improvement

On the basis of the review and evaluation of the program, the evaluation committee may provide recommendations for improving the relevance of the NIOSH research program to health and safety conditions in the workplace and the impact of the research program on health and safety in the workplace.

Appendix A Framework Document
Appendix B Methods and Information Gathering
Appendix C List of NIOSH and Related Materials Collected in the Process of the Evaluation

tp://www.nap.edu/catalog/12639.html

REFERENCES

IOM and NRC (Institute of Medicine and National Research Council). 2006. *Hearing loss research at NIOSH.* Committee to Review the NIOSH Hearing Loss Research Program. Rpt. No. 1, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

IOM and NRC. 2008. *The personal protective technology program at NIOSH.* Committee to Review the NIOSH Personal Protective Technology Program. Rpt. No. 5, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

IOM and NRC. 2009. *Traumatic injury research at NIOSH.* Committee to Review the NIOSH Traumatic Injury Research Program. Rpt. No. 6, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2007. *Mining safety and health research at NIOSH.* Committee to Review the NIOSH Mining Safety and Health Research Program. Rpt. No. 2, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2008a. *Agriculture, forestry, and fishing research at NIOSH*. Committee to Review the NIOSH Agriculture, Forestry, and Fishing Research Program. Rpt. No. 3, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2008b. *Respiratory diseases research at NIOSH*. Committee to Review the NIOSH Respiratory Diseases Research Program. Rpt. No. 4, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2009a. *The health hazard evaluation program at NIOSH.* Committee to Review the NIOSH Health Hazard Evaluation Program. Rpt. No. 7, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2009b. *Construction research at NIOSH*. Committee to Review the NIOSH Construction Research Program. Rpt. No. 8, Reviews of Research Programs of the National Institute for Occupa-

tp://www.nap.edu/catalog/12639.html

tional Safety and Health. Washington, DC: The National Academies Press.

Rutstein, D.D., R. J. Mullan, T. M. Frazier, W. E. Halperin, J. M. Melius, and J. P. Sestito. 1983. Sentinel health events (occupational): A basis for physician recognition and public health surveillance. *American Journal of Public Health* 73(9):1054–1062.

Williams, V. L., E. Eiseman, E. Landree, D. M. Adamson. 2009. *Demonstrating and communicating research impact: Preparing NIOSH programs for external review.* Santa Monica, CA: RAND.

tp://www.nap.edu/catalog/12639.html

4

Improving the Evaluation Process

The evaluation of eight National Institute for Occupational Safety and Health (NIOSH) programs presented challenges in providing a consistent approach to the assessment of a set of diverse research programs, each involving numerous activities. The goal of the series of evaluations was the improvement of the programs and ultimately—given the mission of NIOSH—the improvement of worker safety and health. This chapter considers how the experience gained during this multiphase evaluation effort could be applied to improving future evaluation efforts.

The following sections are based on the framework committee's detailed discussions with those who used the framework (members of the evaluation committees), the sponsor's experience with the evaluation reports, and input from external stakeholders.[1] The resulting revised framework (Chapter 3), and suggestions for its use, are intended as a guide for future evaluations of NIOSH programs and may prove useful for evaluations of research programs in other settings. The chapter is not a summary of the recommendations of the eight program evaluations. Those recommendations and the accompanying detailed evaluations of each program are available in the individual committee reports (IOM and NRC, 2006, 2008, 2009; NRC and IOM, 2007, 2008a,b, 2009a,b).

[1]The chapter draws on the November 24, 2008, workshop organized by the framework committee and held at The National Academies, "Evaluating NIOSH Programs: Lessons Learned and Next Steps." Workshop participants included NIOSH program and senior staff, members of the NIOSH Board on Scientific Counselors, evaluation committee members, framework committee members, and National Academies' staff. Findings also reflect the experiences relayed throughout the eight studies by evaluation committee members and NIOSH staff.

PREPUBLICATION COPY—UNCORRECTED PROOFS

tp://www.nap.edu/catalog/12639.html

EVALUATION FRAMEWORK

Applying the Logic Model and Assessing Outcomes

As discussed in Chapter 2, the logic model is used increasingly by program evaluations to delineate what a program does and accomplishes into the categories of inputs, activities, outputs, and outcomes (see Figure 2-1). The focus of this series of NIOSH program evaluations was on assessing the relevance and impact of the NIOSH program on reducing work-related injuries, illnesses, and hazardous exposures (end outcomes). Initially a number of framework committee members had concerns about developing evaluation criteria that hold a research agency accountable for end outcomes when so many of the factors that affect end outcomes (e.g., workplace conditions, regulations) are outside of the agency's authority or control. Many members of the evaluation committees shared similar concerns.

Several aspects of the framework that were ultimately adopted made the focus on end outcomes more acceptable. The first was the fact that the framework calls for the identification of external factors and provides a fairly extensive list of examples. This provides an opportunity for evaluators to identify obstacles to or promoters of research impact and underscores the changes necessary outside the research agency to improve worker safety and health. A more conventional approach to research evaluation that focuses on knowledge generation would not typically allow for making explicit statements about these obstacles. The second aspect was the clear direction that intermediate outcomes could be used as measures of success in the absence of end outcomes. This was particularly appropriate for exposure-illness relationships that have years of latency before diagnosis. Intermediate outcomes are observable results that could plausibly lead from outputs to end outcomes. Intermediate outcomes (e.g., policy change, adoption of NIOSH-developed technologies in the workplace) may serve as the short- and medium-term proxies for expected end outcomes.

Setting the metric for program success at demonstrating an impact on end outcomes is laudable. Evaluation and framework committee members give NIOSH a great deal of credit for holding their research programs accountable for real outcomes that affect life and health. Evaluations of research supported by other federal agencies often focus on output productivity and intermediate outcomes and do not hold the agency accountable for real-world impacts. This may be appropriate for

tp://www.nap.edu/catalog/12639.html

the goals of some research programs, but NIOSH's focus is on applied research and thus the end outcomes are appropriately focused on improving worker safety and health.

The challenge for the evaluation committees was to search for data on end outcomes while also thoroughly examining the multiple pathways leading to intermediate outcomes and from there to end outcomes. Particularly in the absence of good end outcome data, a frequent finding, attention was often focused on the most observable pathways between program activities and end outcomes. Additional consideration in future evaluations could be given to important, but often less obvious, intermediate outcomes that include:

- Reframing of problems and issues in workplace safety and health in ways that are absorbed into general discussions;
- Contributing indirect knowledge—NIOSH research results that build a body of knowledge that as a whole influences regulation or employer action;
- Accumulating expertise, among NIOSH personnel and in the broader community of NIOSH partners, to allow quick response as needs arise (i.e., building human capital); and
- Building relationships among institutions that need to work together to improve end outcomes (i.e., building social capital).

Measures of these types of contributions are seldom readily available as they are often difficult to quantify and assess. The framework committee views reports from external stakeholders as the most accessible source of this information. In future evaluations, the questions posed by evaluation committees to relevant stakeholders could include this larger set of intermediate outcomes.

A specific issue discussed during the workshop was whether information and technology transfer-activities should be considered in the evaluation and scoring of the program's relevance (as in the initial evaluation framework) or of the program's impact. Research transfer or translation efforts encompass NIOSH program activities conducted to increase the likelihood that research results will impact worker safety and health as well as research aimed at identifying effective translation approaches. External stakeholders frequently use the results of NIOSH research and training to develop programs, processes, materials, or technologies that incorporate research findings in the workplace. Thus, transfer activities are important to the evaluation of both relevance and

tp://www.nap.edu/catalog/12639.html

impact. Because the evaluation framework was designed to assess the efforts of the NIOSH research programs, the framework committee decided to continue to include the evaluation of the extent to which NIOSH programs engage in transfer activities as part of the relevance scoring criteria (Chapter 3).

The framework committee also noted that the logic model approach has some limitations because of its linearity including challenges in handling the multiple inputs, outputs, and external factors and the iterative actions that may contribute toward an outcome being achieved.

Relevance and Impact Ratings

As requested by NIOSH from the outset of this evaluation process, the statement of task included the stipulation that numerical rating systems from 1 to 5 be used to assess the relevance and the impact that the NIOSH program has had on improving worker safety and health. The initial statement of task did not explicitly state that an integer score was needed, but a subsequent revision made this explicit.

Evaluation committee chairs and members indicated that the task of scoring posed significant challenges to the evaluation committees. Initially, many committees struggled with the utility of scoring and were concerned that the restriction to integer scores of 1 to 5 would not convey the nuanced and detailed evaluations they were conducting. Furthermore, some committee members were concerned that a low rating might result in budget reductions or terminated programs. Other committee members thought a high score would not encourage an excellent program to further improve. In several instances scoring took a great deal of the committees' time both at meetings and subsequently through conference calls and emails. Although this was considered a drawback, some members of the evaluation committees indicated that the rating process promoted a detailed prose commentary and in-depth review.

NIOSH staff recognized the challenges that the evaluation committees faced in determining integer-only scores for relevance and impact using a 1-to-5 scale. However, during the November workshop, NIOSH staff reported that the prospect of receiving impact and relevance scores had a motivational effect, in general, on the NIOSH program staff who were assembling the evidence packages and responding to additional requests for information from the evaluation committees.

tp://www.nap.edu/catalog/12639.html

Throughout the evolution of the framework, the scoring criteria have been clarified. The wording for each of the scores for relevance discusses the extent to which the NIOSH program addresses priority research areas and the wording for each of the impact scores discusses the contribution that the NIOSH program has made to worker safety and health. Furthermore, the framework emphasizes that the scoring systems serve as starting points for descriptive text that defines what the scores meant and the program's strengths and limitations that led to the scores. Although there was initial concern that the scores would be the only endpoint noted from the evaluation reports, the committee is pleased to see the detailed strategic plans and action plans developed since publication of the evaluation reports in response to the recommendations.

Integer Scores Versus Decimal Scores

At the outset of the process when developing the scoring criteria, the framework committee had extensive discussions about the nature of the scoring scales and the advantages and disadvantages of using decimals. Although a bigger range of integers or the use of decimals could add finer granularity to the ratings, these options present increased challenges in maintaining consistency across the work of a series of evaluation committees. It was also pointed out that use of decimals implies more accuracy than was possible. In revisiting the scoring issues at the end of the eight evaluations, the framework committee decided to stay with the integer approach using a 5-point scale because it offers greater opportunity for consistency.

Rating Quality

During the workshop, participants discussed the question of whether a specific numerical rating for quality should be added to the task of future NIOSH program evaluations. The issue of developing a process for rating research quality had been discussed early on in the development of the framework, and it was decided that the evaluation of quality was implicit in the evaluation of relevance and impact. Similar thoughts were expressed by workshop participants who stated that the quality of efforts will be seen in the extent and nature of the outcomes. If there is poor quality in the process, a significant intermediate or long-term outcome

tp://www.nap.edu/catalog/12639.html

would not be as likely. Low ratings for program relevance or impact would likely cause program managers to examine research quality as well as other issues such as budget, staffing, or physical plant. Participants noted that the evaluation framework already includes many issues regarding quality. Furthermore, the framework provides opportunities for evaluation committees to include examples of best practices as well to describe the components of an ideal program.

Human Capital Outputs

The critical role of various NIOSH research programs in the training and development of scientists and other professionals was discussed during the workshop and by the framework committee. The framework committee did not see the need to numerically rate this parameter because of lack of data (particularly related to outcomes). The evaluation framework addresses many of the relevant human capital issues through questions about research partners and about training and education as part of transfer activities.

Use of the Framework Document

The general consensus of workshop participants and evaluation committee members was that the evaluation framework was useful but that the extent of close adherence to the details of the framework varied among the evaluation committees. Some evaluation committees adhered rather closely to the framework guidelines and structure, and others used it more as a starting point and reference document. Some evaluation committees thought the structure and format presented by the framework for evaluation was rather rigid, while others found that the framework provided committee members with the latitude needed to individualize the evaluation to fit a specific NIOSH program.

In forming each evaluation committee, one or two members of the framework committee were asked to serve as a liaison or as an evaluation committee member. Additionally, the chair of the framework committee provided an introductory presentation on the framework for each of the evaluation committees. Getting evaluation committee members "buy-in" to follow the framework in conducting the evaluation was often a challenge, and further efforts throughout the process may have improved its use and application. Evaluation committee members differed in the ex-

tp://www.nap.edu/catalog/12639.html

tent to which they used the framework when constructing the narrative. One concern was the limited amount of time allotted by many of the evaluation committees in their initial meetings for discussion of the framework, the evaluation process, and the purpose of the evaluation. Furthermore, the framework committee now recognizes that familiarity and understanding of the framework could be enhanced if the chairs of the evaluation and framework committees discussed the framework and its implementation in depth prior to the first meeting of the evaluation committee.

The two evaluation committees that evaluated programs with major nonresearch components (the Health Hazard Evaluation Program and the Personal Protective Technology Program) found that the framework worked well and with only minor adaptations. These studies had slightly revised statements of task and expanded the rating criteria to include the nonresearch components applicable to each study.

COMPOSITION OF THE EVALUATION COMMITTEES

Producing high-quality and fair assessments begins with the composition and balance of the membership of the evaluation committees. The framework committee and workshop participants believed the independence and credibility of the external organization conducting the evaluations was critical in convincing stakeholders and other interested parties to participate and provide input to the evaluation committees. In following the processes of the National Academies, each evaluation committee member was selected for his or her professional expertise within the breadth of disciplines relevant to the specific program to be evaluated including expertise in program evaluation. Each committee conducted a thorough bias and conflict of interest discussion at its first meeting.

AGENCY INPUTS TO THE EVALUATION

Agency Presentations

At the first meeting of each evaluation committee, NIOSH program and management staff provided briefings on the goals of the evaluation process from the agency's perspective and an overview of the program being evaluated. These presentations were professional, thorough, and

tp://www.nap.edu/catalog/12639.html

detailed. Many evaluation committee members indicated that additional time on the first meeting agenda for discussion with NIOSH staff would have been helpful. Additionally, this initial overview would have benefitted from including more information on how NIOSH sees the program as fitting into the broader context of other research efforts on this topic (see section below on Scope of the Field).

Framework and evaluation committee members expressed interest in hearing more from the NIOSH program staff on what they view as the program's challenges and its opportunities for improvement. Greater emphasis on self-assessment prior to external evaluation could provide valuable information and facilitate a greater understanding of each program.

Evidence Packages

The term *evidence package* refers to the collection of information provided by NIOSH staff to each evaluation committee. Each evidence package consisted of program descriptions, staffing levels, program goals and objectives, and details on the program and its accomplishments. The evidence packages were structured to follow the logic model developed by NIOSH with consultation from RAND Corporation (Williams et al., 2009). Although the evidence package was the core input to each evaluation, it was one of many sources of information that the evaluation committees assessed. Other sources included committee requests to NIOSH for additional information, stakeholder presentations and other input, scientific workshop presentations, and in some cases, site visits.

Although some evidence packages were quite long (more than 900 pages) and detailed, the NIOSH program staff learned throughout the course of the evaluation what evidence was key and what was superfluous to the committee's work. As a result, later evidence packages were more focused. As on any committee, some members wanted more detail and others preferred the big picture. For many committee members, the amount of information was initially overwhelming, but once the organization of information was understood, each member could identify the sections relevant to his or her contributions to the report.

An overview or executive summary of the evidence package is suggested to orient the evaluation committee to program goals, budget, staffing, and congressional mandates as well as to provide highlights of the

tp://www.nap.edu/catalog/12639.html

program's successes and challenges. Several NIOSH evidence packages included such a summary that was helpful to the committee.

Because most evaluation committees divided into working groups that focused on specific subgoals, organizing the evidence package in a similar manner would be particularly useful. Providing an online-searchable version that can be categorized and linked in multiple ways might improve committee use of the package. Committee members noted that the ample references provided in the evidence packages for documentation were quite helpful in assessing the scope and relevance of the program's work.

Evaluation committee members found that the majority of the information was substantive and detailed, but some materials tended to present information from a more positive than critical perspective and the quality of writing in some packages was inconsistent. Some committee members expressed concern that projects that had not been as successful in producing intermediate outcomes were not detailed. Although highlighting success is understandable, it is also important to demonstrate the program's ability to learn from projects that are not successful.

The evidence package was particularly useful in outlining data available on intermediate outcomes but evaluation committee challenges included the lack of surveillance data and the lack of information on end outcomes as related to specific program efforts. Pertinent surveillance data on time trends of the occupational health outcomes relevant to the program's activities should be included in the evidence packages with information, where applicable, about time trends of nonoccupational but relevant health endpoints. For example, assessing the contribution of NIOSH to the decline in work-related homicide is difficult without an assessment of changes in the overall homicide rate over time. Limitations of surveillance data should also be discussed. Because many of the programs rely on the same data sources, such as the Survey of Occupational Injuries and Illnesses, it may be efficient in preparing for future evaluations to develop a generic discussion of the strengths and limitations of surveillance data sources that could be adapted as needed by the various programs.

NIOSH staff noted at the November workshop that the goal in assembling the evidence package was to provide a manageable set of evidence for the evaluation that reflected all aspects of the NIOSH program without overburdening NIOSH staff and budget. NIOSH staff indicated that the process of compiling the evidence packages, although time consuming and intense, was a valuable effort because it focused each pro-

tp://www.nap.edu/catalog/12639.html

gram on its objectives and provided an opportunity to gather internal knowledge and experience to reflect on the program's goals and objectives. Assembling the evidence packages offered NIOSH staff the opportunity for long-term retrospective assessment. Now that the packages have been assembled for the eight NIOSH programs, these documents provide the programs with a baseline and benchmark for future evaluations and a valuable compilation for responding to a wide range of requests for program information.

Quantity and Quality of the Materials

As noted above, the evaluation committees found a great deal of useful information in the evidence packages. During the November workshop, breakout session participants discussed the need for NIOSH staff to be more consistent in categorizing a program's efforts according to the components of the logic model, particularly in differentiating activities and intermediate outcomes. Operational definitions and examples that could improve the categorization of components consistently across programs might be helpful for NIOSH staff in preparing the evidence packages. An online resource that provides this information could be a useful tool and potentially could be incorporated as an integral part of NIOSH's workflow, hopefully reducing the time and staffing burden posed by assembling an evidence package.

Budget and Staffing Information

Several evaluation committee members stated that greater clarification was needed on budgetary and staffing processes (e.g., recruitment of new staff). This information was often the topic of follow-up questions by the evaluation committees to NIOSH. These issues were particularly important for evaluating the programs that function through a matrix management approach, such as the Hearing Loss Research Program, where budgetary authority and line items were in multiple NIOSH divisions.

Scope of the Field and Intramural/Extramural Research

The evaluation committees spent considerable time in grappling with how the NIOSH program fit into the larger scope of research efforts

tp://www.nap.edu/catalog/12639.html

funded by the National Institutes of Health (NIH) and other federal agencies, private-sector companies, foundations, and other organizations. For example, in the respiratory diseases area, NIOSH-funded efforts examining the mechanisms of respiratory cancers were questioned in light of extensive research at the NIH. More information would be helpful in the evidence packages on how NIOSH sees its role and niche in context with other research efforts in that area.

The evidence packages varied in the extent to which information was provided on relevant extramural research, particularly with respect to the types of extramural projects, the levels of extramural funding, and the extent to which the intramural program had input on topics for extramural funding. Many of the evidence packages provided little data on extramural projects, and the committees therefore varied in the extent to which they addressed extramural research. Additionally, the information provided in the evidence packages on the program's strategic planning efforts did not include much attention to the balance or interaction between the extramural and intramural research portfolios related to the program. Understandably the NIOSH research programs must appropriately attend to relevant congressional mandates. Knowledge about such mandates would help evaluation committees gain a better appreciation of how the extramural community's efforts are accounted for in arriving at the intramural research strategy. The evaluation committees also need details on how the intramural research program is engaged in providing input into the development of program announcements or requests for applications for extramural research. Three of the eight evaluation reports specifically recommended greater collaboration and communication between the intramural and extramural aspects of the NIOSH program (IOM and NRC, 2006, 2008; NRC and IOM, 2008b).

Research Prioritizing and Planning Process

Evaluation committee members expressed the need for evidence packages to include more complete discussions of the prioritization of research agendas and the planning processes used by the NIOSH program to establish priorities. Evaluation committees were interested in additional information on how new research programs are initiated and on NIOSH processes for ending efforts that are found to be unproductive. This includes more information on how the NIOSH program identifies emerging research issues. Furthermore, the evaluation committees would

tp://www.nap.edu/catalog/12639.html

have appreciated more information in the evidence packages on processes and balancing of investigator-initiated research versus strategically-planned research.

Overlap Between Programs

Another challenge for NIOSH and the evaluation committees was how to deal with the overlap in areas of research between programs. For example, research on hearing protective devices is relevant to the NIOSH programs on hearing loss; agriculture, forestry, and fishing; personal protective technology; mining; and construction. However, because of the overlap among programs this research area was not thoroughly addressed by any of the committees, and, therefore, may have been incompletely evaluated. The NIOSH respiratory diseases and mining programs also have overlapping research areas, as is also seen in the traumatic injury and construction areas. Evaluation committees generally relied on the extent to which the issues were discussed by NIOSH staff and presented in the evidence packages.

In addressing the issue of how to present research that overlaps multiple programs, it may be helpful for NIOSH to use a consistently defined approach, such as presenting the information in the evidence package for the program that has the lead on the topic, rather than in all areas where the issue relates. Another possibility is that these overlap areas could be described once, then used by different groups preparing evidence packages for different program areas and adapted as needed. This strategy would increase consistency and avoid duplication of staff effort.

Evidence Packages: Conclusions and Next Steps

After reviewing and considering all the comments regarding the evidence packages, the framework committee believes there has been a positive progression and evolution of the evidence packages over the course of the eight studies. Lessons learned in assembling and using the evidence packages for the initial program evaluations were passed along (through comments by NIOSH staff, evaluation committee members, and National Academies' staff), and adaptations in the content and organization of the packages were made along the way. This was particularly

tp://www.nap.edu/catalog/12639.html

helpful for the latter evaluations as NIOSH staff became more aware of the critical materials and the best ways of organizing the data.

As evaluation becomes an integral part of the ongoing work process at NIOSH, the framework committee hopes that systematic data collection and perhaps online tools will be available to collect and analyze data for future evaluations and their evidence packages. Additional points to consider in further efforts to produce evidence packages include:

- Staff training on the components of the logic model;
- Continued sharing of ideas on improving the evidence packages; and
- An executive summary or overview of the evidence package that outlines the basic information for each program (budget and staffing information, list of major goals, highlights of program successes and challenges). A summary that serves as the introductory section of the evidence package could also provide a guide to the materials and a map of the contents.

The framework committee acknowledges the huge effort made by NIOSH staff in assembling the evidence packages and hopes that the institutional commitment that NIOSH made to this evaluation effort is reflected in NIOSH staff evaluations that value the time and effort that was spent in this endeavor.

Site Visits

Evaluation committees that had the opportunity to make site visits to NIOSH facilities found these visits quite informative to the evaluation process. Site visits provided greater context for the evaluation and allowed committee members an opportunity to see the research and program facilities and interact directly with more program staff. Committee members noted that seeing the day-to-day working structure of the program was worth the additional investment of energy and resources. However, given the distributed nature of some programs with many locations and no one central facility, it was not practical for some evaluation committees to conduct a site visit.

The site visits highlighted the importance to the evaluation committees of informal interactions with NIOSH staff members including op-

tp://www.nap.edu/catalog/12639.html

portunities to observe program staff in their work environments, ask questions, and discuss issues outside of a formal presentation setting.

STAKEHOLDER INPUT

Input from external stakeholders was vital in the evaluation process, particularly given the lack of surveillance and other end-outcome data related to determining the program's impact on reducing hazardous exposures or reducing worker injuries or illnesses. Stakeholders may include individuals from labor unions and other worker groups; employers; manufacturers; health and safety and other professional organizations; federal, state, and local regulatory and public health agencies; certifying organizations; the medical community; academic institutions and groups; agricultural, marine, and other nontraditional industry groups; and the international counterparts to NIOSH and the aforementioned groups.

In accordance with the policies of The National Academies for the formation of balanced committees, all of the evaluation committees included individuals from labor organizations, industry, and other stakeholder groups among its subject-matter experts. The insights of such members during committee deliberations were valuable, but cannot be considered to represent all stakeholders effectively. Therefore, the evaluation framework included instructions to the evaluation committees to seek additional stakeholder input. A significant issue the committees faced, however, was how best to use the limited available time and resources to obtain effective and meaningful stakeholder input.

Although the evaluation committees lacked the time and resources to conduct a rigorous scientific survey of external stakeholders, well-balanced input was sought through active mechanisms, such as presentations at committee meetings, or through passive mechanisms, such as requests for written input disseminated via the website or email. Evaluation committees identified stakeholders through lists provided by the NIOSH programs, through recommendations from committee members, and through open calls for input. Some evaluation committees provided the opportunity for stakeholders to submit input anonymously through online questionnaires.

Discussions on stakeholder input during the November 2008 workshop indicated that although all agreed that meeting time devoted to stakeholder input was quite valuable, workshop participants had a mixed reaction regarding the value of seeking comments through a website or

tp://www.nap.edu/catalog/12639.html

online mailing. This approach offered another avenue for receiving comments, but some committees received few meaningful responses in this way. Explicitly clarifying the nature of information being sought from stakeholders and widely disseminating the availability of opportunities for comment may improve the utility of seeking input online.

The adequacy of these approaches must be decided on independently by each evaluation committee. Flexibility and creativity in obtaining stakeholder input are necessary because no single, prescribed method will be suitable for all evaluations. In addition to the methods described above, future committees may consider obtaining input through the National Occupational Research Agenda (NORA) Liaison Committee, which has worked since 1995 to seek regular input from different types of stakeholder groups. Focus groups of external stakeholders also could provide important insights.

Evaluation committees identified several issues in determining how best to use stakeholder input. Although stakeholders often have an inherent biases, their suggestions regarding current and needed NIOSH research efforts can bring the realities of the workplace into focus. Furthermore, stakeholders can relay accounts of responses to NIOSH program activities—the necessary intermediate outcomes that could ultimately lead to end outcomes.

TIME LINES FOR EVALUATIONS

The contract between NIOSH and The National Academies specified that each evaluation report was to be delivered to NIOSH nine months after the first evaluation committee meeting. Two of the eight committees were able to meet that deadline, another two reports were released within 12 months, and the remaining four reports took 15 to 17 months (Table 1-1).

The discussion during one of the breakout sessions of the November workshop focused on the realities of the 9-month time line and options for future studies. Participants suggested that nine months and three meetings did not provide enough time, particularly for the full evaluation of large-scale programs. Reviewing the evidence package and receiving stakeholder input took considerable amounts of time at the first two evaluation committee meetings. The framework committee proposes that a 12- to 14-month time frame with four meetings would be preferable so that the evaluation committee could have more time to gather informa-

tp://www.nap.edu/catalog/12639.html

tion, carefully consider the body of evidence as a whole, determine the ratings, and develop its recommendations. A longer time line would allow time for an additional meeting. Consideration could be given to holding one of the meetings, where feasible, in proximity to the site where the majority of the research program resides, allowing for a better understanding of the context and a more relaxed discussion with research staff about their priority concerns for success. Smaller programs could potentially be evaluated in the 9-month time frame.

The framework committee understands and appreciates the problems that NIOSH encountered as a result of the delays in completing some of the reports, such as the impact on the budget planning cycle. Given the number of variables (including time to assemble the committee, scheduling committee meetings and site visits, delivery of the evidence packages, and the time line for the peer-review process) and the complexity and depth of analysis required, experience has shown that allowing more time for the conduct of future evaluations would be prudent.

EVALUATION COMMITTEE REPORTS AND RECOMMENDATIONS

Report Format

NIOSH staff members expressed an interest in a more consistent format for the evaluation reports. The framework committee noted that the evaluation framework provides a suggested table of contents for evaluation reports, but leaves decisions on the organization and format to the discretion of the individual evaluation committees. The framework committee believes any efforts to further prescribe the report format might jeopardize the creativity and focus of the evaluation committees. Part of the advance training for evaluation committee chairs might include attention to the report format and the recommended structure.

Evaluation Committee Recommendations

All of the evaluation committee reports included recommendations to NIOSH pertinent to improving the relevance and impact of the program. The recommendations addressed a broad array of issues, including enhanced surveillance and monitoring; strategic planning and goal set-

tp://www.nap.edu/catalog/12639.html

ting; improvements in NIOSH internal and external collaborations; increased integration of extramural and intramural research priority setting; and the need for further emphasis on research-to-practice efforts. Although most of the recommendations were of a general nature, some were highly specific to the individual program.

At the November 2008 workshop, NIOSH staff expressed the challenges they face in addressing recommendations that emphasize increased research efforts without accompanying suggestions on what activities should be curtailed or stopped. The "do-more" recommendations far outnumbered the "do-less" recommendations. Although NIOSH staff appreciated the confidence in the programs implied by these recommendations, they also expressed interest in receiving more guidance from the evaluation committees on where and how to set priorities, particularly if faced with static or declining resources. A suggestion was made that future evaluation committees provide recommendations from both perspectives (increases in resources and flat or declining resources). Alternatively, ranking or providing a priority order of the recommendations was also suggested to assist program managers.

After consideration of these issues, the framework committee suggests that if only a few recommendations are made (e.g., fewer than five) that is, in itself, a de facto prioritization. By choosing a few specific recommendations, as opposed to offering numerous recommendations, it is likely that all of the evaluation committee's recommendations are of major importance. If the report contains a large number of recommendations (more than 10), then the evaluation committee should make an effort to discuss priorities. The discussion in the supporting text should emphasize and fully describe the committee's thoughts on what needs to occur first. In some cases, recommendations are related and time ordered, and so prioritization must reflect these issues.

Like most organizations, NIOSH has limited resources. Whenever possible, the evaluation committees should try to consider the optimal allocation of those finite resources within a specific program when making recommendations. Although funding and cost must always be considered, the evaluation committees should not allow funding constraints to be so limiting that they restrict the committee's freedom and the breadth of its collective thinking in addressing its task. By logical extension, such restrictions could curtail the usefulness of its final recommendations. In providing constructive criticism of the agency and vision for its progress, the evaluation committee should not be constrained to treat funding limitations as if they are immutable.

tp://www.nap.edu/catalog/12639.html

Similarly, the framework committee thought that decisions about prioritizing and categorizing the recommendations are best determined by the individual evaluation committees as the recommendations may be organized differently depending on the program's specific organizational structure and programmatic needs.

The framework committee also emphasizes the constraints of evaluating a set of individual programs. Pursuing the "top" priority recommendation for each individual program might not be the optimal approach to achieve overall NIOSH objectives. The prioritization of recommendations for the entire organization requires additional information and further analysis not available to or appropriate for the individual program evaluations.

SUMMARY

The evaluation of eight NIOSH programs within four years was made as consistent as possible by the initial establishment of an evaluation framework. Having a common statement of task and evaluation approach in the use of the logic model provided a starting point for each evaluation committee. Furthermore, knowing that the endpoint involved the application of defined scoring criteria for rating the program's relevance and impact on reducing work-related injuries, illnesses, or hazardous exposures provided a common definitive goal for the evaluation committees. As with any multiphase effort involving numerous agency staff and multiple committees, there were variations in how the task was completed.

NIOSH is to be commended on its focus on end outcomes and for holding its research and program staff accountable for progress toward those outcomes. Increased opportunities, particularly informal opportunities, for NIOSH staff to discuss issues with evaluation committees would be helpful, as would increased attention by evaluation committees to some of the more indirect measures of intermediate outcomes.

This chapter has addressed both the framework and its application. The evaluation committee's experiences with the use of the framework document were variable. A number of areas in which the evaluation process could be improved were highlighted, including the provision of more information in the evidence packages on priority setting efforts and on budget and staffing; further input on extramural research and the connections between the intramural and extramural program objectives; en-

tp://www.nap.edu/catalog/12639.html

suring plenty of opportunities for input by external stakeholders and agency staff; considerations regarding the time line for the evaluations; and options for other types of recommendations or ways of categorizing recommendations in the evaluation reports. The framework committee has made some changes to the evaluation framework (Chapter 3) based on this assessment of lessons learned, and as indicated in the following chapter has thoughts on future evaluation efforts.

REFERENCES

IOM and NRC (Institute of Medicine and National Research Council). 2006. *Hearing loss research at NIOSH.* Committee to Review the NIOSH Hearing Loss Research Program. Rpt. No. 1, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

IOM and NRC. 2008. *The personal protective technology program at NIOSH.* Committee to Review the NIOSH Personal Protective Technology Program. Rpt. No. 5, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

IOM and NRC. 2009. *Traumatic injury research at NIOSH.* Committee to Review the NIOSH Traumatic Injury Research Program. Rpt. No. 6, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2007. *Mining safety and health research at NIOSH.* Committee to Review the NIOSH Mining Safety and Health Research Program. Rpt. No. 2, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2008a. *Agriculture, forestry, and fishing research at NIOSH.* Committee to Review the NIOSH Agriculture, Forestry, and Fishing Research Program. Rpt. No. 3, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2008b. *Respiratory diseases research at NIOSH.* Committee to Review the NIOSH Respiratory Diseases Research Program. Rpt. No. 4, Reviews of Research Programs of the National

tp://www.nap.edu/catalog/12639.html

Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2009a. *The health hazard evaluation program at NIOSH.* Committee to Review the NIOSH Health Hazard Evaluation Program. Rpt. No. 7, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2009b. *Construction research at NIOSH.* Committee to Review the NIOSH Construction Research Program. Rpt. No. 8, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

Williams, V. L., E. Eiseman, E. Landree, and D. M. Adamson. 2009. *Demonstrating and communicating research impact: Preparing NIOSH programs for external review*. Santa Monica, CA: RAND.

tp://www.nap.edu/catalog/12639.html

5

Recommendations for Moving Forward

Drawing on the lessons learned in developing the evaluation framework and in applying the framework during eight evaluations, this chapter provides the framework committee's recommendations on next steps for National Institute for Occupational Safety and Health (NIOSH) program evaluations, particularly from a long-range perspective. These recommendations may also be informative for other federal agency program evaluations.

ONGOING EVALUATION

Evaluation of research programs at regular intervals has become the norm, with the trend toward internal management reviews supplemented by periodic evaluation by external parties (Chapter 2). Competently done external evaluation removes the unconscious bias of managers with regard to their programs; takes organizational competition out of the assessment; and usually provides new insights while reinforcing some of what managers already knew, but could not act on.

Findings of program evaluations serve a variety of purposes. Direct and instrumental uses of the evaluation report include specific modifications to the program through actions such as increasing or decreasing its budget, redefining its program objectives, or more obviously tying its work to the work being done in sister programs. Evaluation reports can also be used by program staff to call attention to research problems or needs, thus raising the salience of an issue within the agency. Often, use occurs that is of a longer range and more diffuse nature. Agency staff and

tp://www.nap.edu/catalog/12639.html

researchers may, by virtue of reading an evaluation report, gain an appreciation for the external factors that have shaped a research agenda, or they may see common threads across a set of evaluations that identify the need for new agency-wide strategic objectives. These less direct, longer term uses of evaluation products can function as a way of gradually infusing an organization with new ideas from the outside to the inside. Furthermore, external stakeholders or advocacy groups can use an evaluation report to call attention to needed research and priorities.

The committee commends NIOSH for undertaking the recent external review of a series of their programs by the National Academies, and in particular for requesting that the focus of the evaluations be on the relevance and impact in reducing work-related illnesses, injuries, or hazardous exposures. Many evaluations, particularly of research programs, stop with assessments of outputs, such as the number of peer-reviewed publications, and do not take into account the research transfer steps, the external factors that influence program activities and outcomes, or the need to use the program's impact on intermediate or end outcomes as the metric for a successful program.

In an effort to continue the forward momentum, the committee recommends that NIOSH establish a system for periodic external evaluation of its programs. The system does not need to be modeled on the evaluations just carried out by the National Academies; indeed, many options are available, and NIOSH may want to incorporate elements of several in its overall approach. For example, the following types of evaluation approaches may be considered:

- From time to time, an outside look at agency-wide processes is necessary.
- Regular external review at the broad program level should be complemented with internal self-assessments on a more regular basis.
- Formative evaluations at the program design or implementation stages can be useful for new programs.
- Special studies of the long-term impacts of the program (such as the development of human capital) could be undertaken.

Whatever the composition or structure of external review, the research program should conduct a self-study prior to external review. This is analogous to the self-study that an academic program would conduct prior to accreditation. Much of the evaluation framework presented in

tp://www.nap.edu/catalog/12639.html

Chapter 3, including the scoring system, may be useful in self-evaluations.

These different approaches might emphasize different criteria for NIOSH effectiveness, balancing the appropriate focus on long-term impact in the workplace with attention to other public benefits. As noted above, in addition to research program evaluations, external evaluations of agency-wide initiatives or processes can be useful. The agency is more than the sum of its parts, and a broad review can provide perspective not gained in a cross-section of assessments of individual programs. Further, an emphasis on how the program contributes to the overall NIOSH program could be added, as the current framework is focused on issues specific to the program and its goals.

Recommendation 1 ***Continue Systematic External Evaluations***
NIOSH should establish a system for periodic external evaluation complemented by internal self-assessments on a regular basis. Program or agency-wide evaluations should begin with strong self-evaluation efforts that allow the program or agency to assemble and analyze data and act on relevant findings concerning the program's strengths, weaknesses, and opportunities.

CONTINUE TO BOLSTER RESEARCH TRANSLATION EFFORTS

From the perspective of NIOSH, research transfer or translation can be seen to encompass activities that staff and researchers engage in to increase the likelihood that the results of research will be used to improve worker safety and health as well as studies of research translation conducted or funded to increase knowledge about which approaches are most effective. Research transfer is a commendable new emphasis in the agency, and one that the evaluation committees noted has provided a number of positive intermediate outcomes.

NIOSH refers to research translation as "r2p"—Research to Practice. However, the committee noted that the focus of NIOSH's efforts on research-to-practice tends to be on interventions, demonstration projects, and control technology rather than the results of health effects research. NIOSH should expand its portfolio of r2p efforts to formulate an approach to translation of health effects research. The research observation

tp://www.nap.edu/catalog/12639.html

of a new or more fully understood association between exposures and illness can lead to changes in use of existing technology or control programs and requires research translation efforts at a level similar to the development and application of new technology in the workplace to reduce exposure.

Much remains to be learned about how to improve the likelihood that research translation efforts (e.g., site visits, demonstrations, partnerships) and products (e.g., websites, newsletters, CD-ROMs, training materials) will positively impact worksites. The systematic gathering and interpretation of early-stage feedback about research-to-practice processes and products prior to release or deployment is a form of research that focuses on the intended user, the user's work context, and how improvements can be made to be most compatible with current work practices and conditions. These types of preintervention research efforts can be conducted with a small number of potential adopters and at little cost as long as respondents are somewhat representative of the larger worksite population that will be targeted. Feedback can then be used to identify which information channels are preferred by the target audience, which messages and images best communicate the positive attributes of the research results in question, which types of spokespeople are most effective at producing interest and inquiries by potential adopters, and how NIOSH innovations themselves can be redesigned to be more compatible with real-life workplace conditions and constraints. Conducting formative studies such as these can come to constitute a key NIOSH r2p skill set, although capacity building would be warranted. Further development of these skills within NIOSH may become an example of how social science can contribute to the agency's worker protection and workplace safety mission.

NIOSH is also warranted in soliciting and funding behavioral research about the dissemination of safety and occupational health outputs and intervention programs, the barriers and facilitators that affect adoption of effective practices and processes in the workplace, and the study of effective implementation of NIOSH research results within organizations. The evaluation committee that assessed the NIOSH Agriculture, Forestry, and Fishing (AFF) Research Program made the following recommendation: "NIOSH should conduct research on the science of knowledge diffusion to identify effective methods for AFF research-to-practice programs" (NRC and IOM, 2008a, p. 12).

Continued contributions by NIOSH to research on improving the effectiveness of translation efforts will ensure the consideration of the dy-

tp://www.nap.edu/catalog/12639.html

namics that characterize workplace safety and occupational health. The committee believes NIOSH has a role to play not just in demonstrating and testing research-to-practice approaches, but also in documenting and testing its inverse, practice-to-research. Often, the most effective research translation occurs through iterative learning. Practitioners can learn from researchers, but it is at least as important for researchers to learn from practitioners so that the new knowledge, practices, programs, and technologies that researchers create are informed by real-world workplace conditions.

Recommendation 2 ***Continue to Build and Improve Research Translation Efforts***
NIOSH should continue to build and improve its research translation efforts with an emphasis on:

- **ongoing assessment and improvement of its research translation efforts through formative evaluation processes of listening to those in the workplace (workers and employers) and beyond (product designers, architects, health care providers, etc.), both to identify intervention needs and to provide early feedback regarding research translation products to improve the interventions; and**
- **building the capacity to implement and evaluate research translation efforts, both as research-to-practice and as practice-to-research.**

ENHANCE OCCUPATIONAL HEALTH AND SAFETY SURVEILLANCE

The logic model approach to evaluation of NIOSH programs—used as the basis for the framework presented in this report—relies heavily on surveillance data on health outcomes and workplace exposures to evaluate strategic priorities and assess program impact. Surveillance data are also critical program inputs, and the extent to which research programs have considered surveillance data in setting research priorities is an important determinant of program relevance. The importance of surveillance extends far beyond research; it is also critical for effectively targeting and evaluating intervention activities at the national, state, and

tp://www.nap.edu/catalog/12639.html

local levels and for the strategic planning needed to develop, implement, and assess these interventions. Although a comprehensive system for tracking fatal occupational injuries in the United States is in place, the current approaches to surveillance of occupational illnesses and nonfatal occupational injuries are fragmented and incomplete (Azaroff et al., 2002), and only limited surveillance data on exposure to hazards are available. All eight evaluation committee reviews of NIOSH programs completed to date identified the need for better surveillance, and seven of the eight evaluation reports included specific recommendations calling for improved surveillance and additional surveillance research (IOM and NRC, 2006, 2008, 2009; NRC and IOM, 2007, 2008a,b, 2009a).

NIOSH includes surveillance in its mission and is engaged in a number of surveillance activities, both intramural and in collaboration with state partners, using a variety of data sources. Although NIOSH has made many important contributions to surveillance, the committee that evaluated the NIOSH Traumatic Injury Research Program noted that "these projects do not appear to be a part of a coordinated interagency strategy to improve national surveillance of traumatic nonfatal occupational injuries" (IOM and NRC, 2009, p. 52).

The Survey of Occupational Injuries and Illnesses conducted by the Bureau of Labor Statistics (BLS), the official source of statistics on nonfatal, work-related injuries and illnesses, is a valuable resource, but has many limitations. This employer-based data source currently excludes approximately 22 percent of the workforce and fails to capture most occupational illnesses (Leigh et al., 2004). Nonfatal injuries may be substantially underreported (Azaroff et al., 2002; Welch et al., 2007; Boden and Ozonoff, 2008), and there is concern that underreporting may vary by type of injury and worker or workplace characteristics. Important information about employer health and safety practices that could identify correlations between employer practices and good health and safety records are not collected. It would be unfortunate if the recommendations for improved surveillance in individual NIOSH program reviews led to fragmented surveillance activities. The identification of the need for improved surveillance across NIOSH research programs underscores the need for a comprehensive and coordinated interagency plan for surveillance of work-related injuries, illnesses, and hazards.

NIOSH, which has epidemiologic capacity and experience working with a wide range of health data sources, could play an important leadership role in coordinating efforts of relevant federal agencies, including but not limited to BLS, the Occupational Safety and Health Administra-

tp://www.nap.edu/catalog/12639.html

tion, Mine Safety and Health Administration, National Center for Health Statistics, and Department of Transportation. States, which have the legal authority to require disease reporting and to collect health data, should be included as integral partners in developing and implementing a comprehensive surveillance plan. Such a plan should go beyond improvements in the existing employer-based data sources to include nonemployer-based data sources such as hospital and other medical data systems and population-based surveys such as the National Health Interview Survey and the Behavioral Risk Factor Surveillance System. The potentially highly important role of electronic health records and new integrated data systems in surveillance efforts should be addressed. The plan should also include a surveillance research agenda and a proposed mechanism for ongoing interagency communication and coordination.

Surveillance is a necessity for monitoring long-term progress in reducing hazardous exposures and work-related injuries and illnesses.

Recommendation 3 ***Increase and Improve Surveillance to Benchmark Progress***

NIOSH should increase and improve surveillance of work-related injuries, illnesses, exposures, and working conditions so that information needed to assess program relevance and impact will be available for future evaluations. Enhanced surveillance should prove informative in balancing research priorities.

INTEGRATE EVALUATIONS OF EXTRAMURAL AND INTRAMURAL RESEARCH

Obtaining the full picture of NIOSH's work in a specific area of research requires examining the relevant intramural and extramural research. However, the evaluation committees found that the extent to which the intramural and extramural components at NIOSH are currently separated makes it difficult to conduct such an assessment. Several of the evaluation committees noted a disconnect between the intramural and extramural programs. For example, the evaluation report on the NIOSH Hearing Loss Research Program stated, "In some cases, however, intramural researchers have not made themselves aware of relevant extramural research, which may have resulted in limited opportunities for effective collaboration" (IOM and NRC, 2006, pp. 122–123). Similarly,

tp://www.nap.edu/catalog/12639.html

the report on the Traumatic Injury Research Program noted in a recommendation, "NIOSH should review its practices and take steps to improve the opportunities for intramural and extramural researchers, including state occupational public health programs, to communicate and collaborate without excessively directing extramural research to the detriment of scientific creativity" (IOM and NRC, 2009, p. 14).

Although the framework committee fully supports external scientific review to determine merit for funding investigator-initiated research, the evaluation committees noted that few avenues are currently available by which NIOSH staff can provide intramural input into the development of priorities for extramural research. Larger research agencies, such as the National Institutes of Health, fund full-time staff members to interface between the intramural and extramural programs, but this may not be a feasible option for NIOSH due to funding constraints. The evaluation report on the Personal Protective Technology (PPT) Program noted that improvements are being made in this area: "PPT Program staff members have also reported increased opportunities for dialogue with the NIOSH Office of Extramural Programs in the past year regarding priorities for funding. The committee urges NIOSH to consider ways in which the PPT Program could have greater input into the extramural priority process at NIOSH and increased participation in drafting requests for grant applications" (IOM and NRC, 2008, p. 122).

Future evaluation efforts need to focus on examining the relationship between intramural and extramural research in strategic planning for a cohesive research program that addresses program goals and the overall NIOSH mission.

Recommendation 4 ***Integrate Evaluations of Intramural and Extramural Research***

Future evaluations should systematically consider intramural and extramural research activities, in terms of both evaluating the impact and relevance of each type of research and assessing the extent to which intramural and extramural research are integrated in strategic planning.

ON THE HORIZON

Over the course of its work during the past four years, the framework committee has seen continuing developments in occupational safety and

tp://www.nap.edu/catalog/12639.html

health that will have significant impacts in the decades ahead. The committee concludes this report with an overview of a few of the issues that may need to be considered in future evaluations of progress made in improving worker health and safety:

- Defining "the workplace": Traditional workplaces and job stability are rapidly changing. For example, some forms of work are increasingly conducted outside the physical workplace. Internet accessibility permits work to occur in homes and other locations through a variety of telecommuting arrangements. In addition, the mobility of some individual workers' jobs prevents traditional approaches to monitoring these workers and their work risks in a reliable fashion. Appropriate surveillance for workplace injuries, illnesses, exposures, and risks consequently has become much more complex and presents many new challenges for assessing and preventing work-related problems.
- Occupational health and public health: The worksite is increasingly recognized as a venue to improve population health. New emphasis is being placed on changes in workplace policies and practices that promote healthy lifestyle choices and help prevent chronic diseases such as diabetes and stroke. As reflected in the NIOSH WorkLife Initiative, new integrated approaches in worksites that address both occupational and nonoccupational risks are needed and will broaden considerations when evaluating occupational health and safety programs.
- Ethics: Ethical issues relevant to occupational health and safety research are expanding beyond the bounds of institutional review boards and include issues of equity in addressing underserved and vulnerable populations and attention to working conditions in the small-business sector.
- Worker demographics: The changing demographics of the workforce will need to be considered in future evaluations. For example, employment in the services sector continues to grow, the U.S. labor force continues to age, and the workforce is becoming more racially and ethnically diverse.
- Economic challenges: Challenging economic times will likely bring as yet unknown changes to the workplace. A forward look at these changes will be necessary to keep pace with the needs for health and safety of the U.S. and global workforce.

tp://www.nap.edu/catalog/12639.html

REFERENCES

Azaroff, L. S., C. Levenstein, and D. H. Wegman. 2002. Occupational injury and illness surveillance: Conceptual filters explain underreporting. *American Journal of Public Health* 92(9):1421–1429.

Boden, L., and A. Ozonoff. 2008. Capture-recapture estimates of nonfatal workplace injuries and illnesses. *Annals of Epidemiology* 18(6):500–506.

IOM and NRC (Institute of Medicine and National Research Council). 2006. *Hearing loss research at NIOSH.* Committee to Review the NIOSH Hearing Loss Research Program. Rpt. No. 1, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

IOM and NRC. 2008. *The personal protective technology program at NIOSH.* Committee to Review the NIOSH Personal Protective Technology Program. Rpt. No. 5, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

IOM and NRC. 2009. *Traumatic injury research at NIOSH.* Committee to Review the NIOSH Traumatic Injury Research Program. Rpt. No. 6, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

Leigh, J. P., J. Marcin, and T. R. Miller. 2004. An estimate of the U.S. government's undercount of non-fatal occupational injuries. *Journal of Occupational and Environmental Medicine* 46(1):10–18.

NRC and IOM. 2007. *Mining safety and health research at NIOSH.* Committee to Review the NIOSH Mining Safety and Health Research Program. Rpt. No. 2, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2008a. *Agriculture, forestry, and fishing research at NIOSH.* Committee to Review the NIOSH Agriculture, Forestry, and Fishing Research Program. Rpt. No. 3, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2008b. *Respiratory diseases research at NIOSH.* Committee to Review the NIOSH Respiratory Diseases Research Program. Rpt. No. 4, Reviews of Research Programs of the National

tp://www.nap.edu/catalog/12639.html

Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2009a. *The health hazard evaluation program at NIOSH.* Committee to Review the NIOSH Health Hazard Evaluation Program. Rpt. No. 7, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

NRC and IOM. 2009b. *Construction research at NIOSH.* Committee to Review the NIOSH Construction Research Program. Rpt. No. 8, Reviews of Research Programs of the National Institute for Occupational Safety and Health. Washington, DC: The National Academies Press.

Welch, L. S., X. Dong, F. Carre, and K. Ringen. 2007. Is the apparent decrease in injury and illness rates in construction the result of changes in reporting? *International Journal of Occupational and Environmental Health* 13(1):39–45.

tp://www.nap.edu/catalog/12639.html

tp://www.nap.edu/catalog/12639.html

A

Meeting Agendas

INSTITUTE OF MEDICINE
NATIONAL RESEARCH COUNCIL

Committee on the Review of NIOSH Research Programs

May 5–6, 2005
Open Session Agenda

May 5, 2005

10:00 a.m.	**Welcome and Introductions** Evan Douple, Ph.D., Scholar, National Academies David Wegman, Ph.D., Committee Chair Lewis Wade, Ph.D., Senior Science Advisor, NIOSH
10:30	**Statement of Charge and Presentations by NIOSH** Lewis Wade, Ph.D., Senior Science Advisor, NIOSH
11:45	**DOE R&D Review** Robert Fri, Senior Fellow Emeritus, Resources for the Future, Inc. Chair, NRC Committee on Perspective Benefits of DOE's Energy Efficiency and Fossil R&D Programs

tp://www.nap.edu/catalog/12639.html

2:30 p.m. **NIOSH Presentation**
W. Gregory Lotz, Ph.D., Associate Director for Science, Division of Applied Research and Technology, NIOSH

3:30 **RAND Presentation**
Victoria A. Greenfield, Ph.D., Senior Economist, RAND

4:30 **Discussion/Revisit Statement of Task**

May 6, 2005

10:30 a.m. **NIOSH Presentation**
Jeffrey L. Kohler, Ph.D. Associate Director for Mining and Construction, NIOSH

INSTITUTE OF MEDICINE
NATIONAL RESEARCH COUNCIL

Committee on the Review of NIOSH Research Programs

July 7–8, 2005
Open Session Agenda

July 7, 2005

9:30 a.m. **Greetings from the Director of NIOSH**
John Howard, M.D., Director NIOSH

9:45 **Presentations on NIOSH Data and Interaction with Other Organizations**
Charles Jeffress, Chief Administrative Officer, Legal Services Corporation (Head of OSHA, 1997–2002)

10:15 Sue Dong, Director, Data Department, Center to Protect Workers' Rights, AFL–CIO

tp://www.nap.edu/catalog/12639.html

10:45	James Weeks, Sc.D., CIH, Senior Scientist, Advanced Technologies and Laboratories International, Inc.
11:15	Mei-Li Lin, Ph.D., Director, Research and Statistical Services, National Safety Council
11:45	Roundtable Discussion

July 8, 2005

8:30 a.m.	**Welcome and Introductions** David Wegman, M.D., Chair
8:45	**Continued Discussion on NIOSH Interaction with Other Organizations** Steven Lerman, M.D., Manager, Occupational and Public Health Division, ExxonMobil

---

INSTITUTE OF MEDICINE
NATIONAL RESEARCH COUNCIL

Committee on the Review of NIOSH Research Programs

April 10–11, 2006
Open Session Agenda

April 10, 2006

10:00 a.m.	**NIOSH Health Hazards Evaluation Program as a Research Program** Allison Tepper, Ph.D., Epidemiologist and Chief of the Hazard Evaluations and Technical Assistance Branch of the NIOSH Division of Surveillance, Hazard Evaluations, and Field Studies
10:45	**Authoritative Recommendations Program as a Research Program** Paul Schulte, Ph.D., Director, NIOSH Education and Information Division

tp://www.nap.edu/catalog/12639.html

11:30 **Discussion with NIOSH Regarding Applicability of Framework to an Evaluation of the HHE Programs**

12:30 p.m. **Continued Discussion with NIOSH**

INSTITUTE OF MEDICINE
NATIONAL RESEARCH COUNCIL

Committee on the Review of NIOSH Research Programs

May 30–31, 2007
Open Session Agenda

May 30, 2007

10:35 a.m. **NIOSH Response to Reports and Framework Document**
Lewis Wade, Ph.D., Senior Science Advisor, NIOSH

Feedback from NIOSH about the Framework Document, the NA review process, and NA Hearing Loss and Mining reports
Feedback from the Framework Committee and NA staff about NIOSH "evidence packages" and presentations to Evaluation Committees
Adapting the Framework to cover nonresearch programs
A potential summative evaluation activity conducted by The National Academies

tp://www.nap.edu/catalog/12639.html

INSTITUTE OF MEDICINE
NATIONAL RESEARCH COUNCIL

Committee on the Review of NIOSH Research Programs

November 24–25, 2008
Open Session Agenda

Workshop on Evaluating NIOSH Programs:
Lessons Learned and Next Steps

November 24, 2008

Workshop Goals:

- Share lessons learned from the NIOSH evaluation process and ideas regarding the framework and the evaluation process.
- Discuss improvements to the evaluation process and the evaluation framework.
- Discuss potential evaluation issues for future studies and their impact on the framework and the evaluation process.

8:00 a.m. **Welcome, Introductions, and Goals for the Workshop**
David Wegman, Chair

8:15 **NIOSH Perspective**
Christine Branche, Acting Director, NIOSH
NIOSH Staff

Discussion

9:00 **NIOSH Efforts in Response to the Reports**
Sarah Felknor, Chair, NIOSH Board of Scientific Counselors
Ray Sinclair, NIOSH

Discussion

9:30 **Panel Discussion: Lessons Learned and Opportunities for Improvement**
Evaluation Committee Chairs (or representative members) and NIOSH Staff

tp://www.nap.edu/catalog/12639.html

9:30	**Evidence Package** Opening Comments, Mark Utell, Chair, Respiratory Disease Report Additional Comments by Panel Members
10:00	**Evidence Process** Opening Comments, John Gallagher, Chair, PPT Report Additional Comments by Panel Members
10:30	**Rating System** Opening Comments, Paul Gunderson, Chair, Agriculture, Forestry, and Fishing Report Additional Comments by Panel Members
11:00	**Other National Academies Models for Program Evaluation** Greg Symmes, Deputy Director, Division on Earth and Life Studies
	Discussion
11:30	**Evaluators' Perspective: Setting the NIOSH Evaluation Framework in Context** Susan Cozzens, Monica Gaughan, Joe Wholey
	Discussion
12:50 p.m.	**Breakout Groups** Breakout Session Goal: Discuss improvements for future evaluations; provide list of suggestions

- **Breakout A: Evidence Package**
  - Tish Davis, facilitator
  - Richard Tucker, rapporteur
- **Breakout B: Evaluation Process (e.g., stakeholder input, site visits, time line)**
  - Jackie Nowell, facilitator
  - Raj Ramani, rapporteur

tp://www.nap.edu/catalog/12639.html

- **Breakout C: Relevance and Impact Rating System**
  - Jim Zuiches, facilitator
  - Fred Mettler, rapporteur
- **Breakout D: Prioritizing and Categorizing Recommendations (zero sum versus additional funds; short-term versus long-term recommendations)**
  - Jorma Rantanen, facilitator
  - Jim Dearing, rapporteur

2:30 **Reports from Breakout Groups and Discussion**

3:15 **Summary**
David Wegman, Chair, Framework Committee

tp://www.nap.edu/catalog/12639.html

tp://www.nap.edu/catalog/12639.html

B

Biographical Sketches of Committee Members

David H. Wegman (*Chair*), is a professor in the School of Health and Environment at the University of Massachusetts–Lowell. He rejoined the faculty in fall 2008 after serving a 5-year term as dean of the School of Health and Environment. Previously he served as chair of the Department of Work Environment, a position he held since its founding in 1987, and he remains an adjunct professor at the Harvard School of Public Health. Prior to 1987 he served as director of the Division of Occupational and Environmental Health at the University of California–Los Angeles (UCLA) School of Public Health and on the faculty at Harvard School of Public Health. Dr. Wegman has focused his research on epidemiologic studies of occupational respiratory disease, musculoskeletal disorders, and cancer. He has published more than 200 articles in the scientific literature. He has also written on public health and policy issues concerning hazard and health surveillance; methods of exposure assessment for epidemiologic studies; development of alternatives to regulation; and use of participatory methods to study occupational health risks. He is coeditor of *Occupational and Environmental Health: Recognizing and Preventing Disease and Injury*, the fifth edition of which was published by Lippincott, Williams, and Wilkins in 2006. His recent work has focused on the examination of health and safety risks among heavy- and highway-construction workers; the study of the relationship of work risks and age among child laborers and older adults; and public health surveillance methods and systems for occupational disease. He has served as chair of the National Research Council (NRC) and Institute of Medicine (IOM) Committees on Health and Safety Needs of Older Workers and the Health and Safety Consequences of Child Labor. He has also been a member of the NRC–IOM Panel on Musculoskeletal Disorders and

PREPUBLICATION COPY—UNCORRECTED PROOFS

tp://www.nap.edu/catalog/12639.html

Work, and the IOM committees to Review the Health Consequences of Service During the Persian Gulf War and to Review Gender Differences in Susceptibility to Environmental Factors. Currently he serves as chair of the NRC–IOM Committee on Review of National Institute for Occupational Safety and Health (NIOSH) Research Programs. He received his B.A. from Swarthmore College and his M.D. and M.Sc. from Harvard University and is board certified in preventive medicine (occupational medicine).

William B. Bunn, III, is vice president of health, safety, security, and productivity at International Truck and Engine Corporation (formerly Navistar International) in Warrenville, IL. Previously, he was medical director and director of health care, workers' compensation, disability, and safety for Navistar International, and prior to that was director of International Medical Services for Mobil Corporation. Dr. Bunn has an appointment as professor of clinical preventive medicine at Northwestern University School of Medicine. He received the Occupational and Environmental Education Foundation Award in 2003, the William S. Knudsen Award in 2002, and the Institute for Health and Productivity Management Corporate Health and Productivity Award in 2001. He chaired the NRC Committee on Department of Energy Radiation Epidemiological Research Programs, and has served on numerous advisory committees, including the Science Advisory Board of the Environmental Protection Agency, Board of Scientific Counselors of NIOSH, and Committee on Clinical Services. He is also a fellow board member and former officer of the American College of Occupational and Environmental Medicine. He received a J.D. and M.D. from Duke University and an M.P.H. from the University of North Carolina.

Carlos A. Camargo Jr. is an associate professor of medicine and epidemiology at Harvard Medical School, an emergency physician at Massachusetts General Hospital and a research epidemiologist at the Channing Laboratory, Brigham and Women's Hospital in Boston, MA. His research focuses on asthma and other respiratory/allergy problems in several large national cohorts (e.g., the Nurses' Health Studies). He also chairs the Steering Committee of the Emergency Medicine Network (EMNet), a research collaboration involving 204 emergency departments. This network has completed numerous multicenter studies and randomized trials focusing on respiratory/allergy emergencies and public health issues. Dr. Camargo is past president of the American College of

tp://www.nap.edu/catalog/12639.html

Epidemiology and served on the 2005 U.S. Dietary Guidelines Advisory Committee and the National Institutes of Health's National Asthma Education and Prevention Program's Third Expert Panel (the group writing the national asthma guidelines). He has authored more than 350 peer-reviewed publications. Dr. Camargo received his M.D. from the University of California–San Francisco; his M.P.H. from the University of California–Berkeley; and his Dr.P.H. from the Harvard School of Public Health.

Susan E. Cozzens is a professor of public policy at the Georgia Institute of Technology, director of its Technology Policy and Assessment Center, and associate dean for research at its Ivan Allen College. She is currently working on research in the fields of science, technology, and inequalities; she continues to work internationally on developing methods for research assessment, as well as science and technology indicators. Prior to joining the faculty at Georgia Tech, she was the director of the Office of Policy Support at the National Science Foundation (NSF). Dr. Cozzens has served as a consultant to numerous organizations, including the Office of Science and Technology Policy, NSF, Office of Technology Assessment, General Accountability Office, National Cancer Institute, National Institute on Aging, and other NIH institutes. She has served on several NRC and IOM committees, including Evaluation of the Sea Grant Program Review Process, Assessment of Centers of Excellence Programs at NIH, Research Standards and Practices to Prevent the Destructive Application of Biotechnology, and the Committee to Review the NIOSH Hearing Loss Research Program. Dr. Cozzens is the past editor of *Science, Technology, & Human Values* and the *Journal of the Society for Social Studies of Science*. She currently is the coeditor of *Research Evaluation*. She earned her Ph.D. in sociology from Columbia University.

Letitia Davis is director of the Occupational Health Surveillance Program in the Massachusetts Department of Public Health, where she has worked for more than 20 years to develop state-based surveillance systems for work-related illnesses and injuries. She has overseen the formation of a physician reporting system for occupational disease, the Massachusetts Occupational Lead Registry, a comprehensive surveillance system for fatal occupational injuries, the Massachusetts Sharps Injury Surveillance System, and a model surveillance system for work-related injuries to children and adolescents younger than 18. She has conducted numerous surveillance research studies exploring use of existing public health data sources

tp://www.nap.edu/catalog/12639.html

to document work-related injuries and illnesses. She is currently engaged in a project using community health center data to document occupational health needs of underserved worker populations. She is also responsible for the development of prevention programs to address identified occupational health problems and advises the department leadership on matters of occupational health policy. Dr. Davis serves as adjunct faculty of the Department of Work Environment at the University of Massachusetts–Lowell and as a visiting lecturer on Occupational Health at the Harvard School of Public Health. She is also a lead consultant in occupational health to the Council of State and Territorial Epidemiologists and has played a leadership role nationally in the effort integrate occupational health into public health practice at the state level. She is a past member of the Board of Scientific Counselors of NIOSH and of the National Advisory Committee on Occupational Safety and Health. Dr. Davis received her doctorate in occupational health from the Harvard School of Public Health.

James W. Dearing is senior scientist at the Institute for Health Research with Kaiser Permanente Colorado, where he directs the Cancer Communication Research Center and codirects the Center for Health Dissemination and Implementation Research (with Russell Glasgow). Until 2006, he was professor and director of graduate studies for the School of Communication Studies at Ohio University and has been a faculty member at Michigan State University, a visiting faculty member at the University of Michigan, and a visiting scholar at the University of California–Berkeley. Dr. Dearing studied under and collaborated with Everett M. Rogers for 20 years. Dr. Dearing's primary area of expertise is the application of diffusion of innovation concepts to challenges of moving evidence-based practices, programs, and policies into practice. He has led research projects about community-based health system reform, mass media agenda setting, community health promotion planning, interorganizational networks, and organizational change. Dr. Dearing has most recently conducted studies sponsored by the NSF; John D. and Catherine T. MacArthur Foundation; The Robert Wood Johnson Foundation; National Heart, Lung, and Blood Institute; and the National Cancer Institute. His most recent book, *Communication of Innovations*, was coedited with Arvind Singhal. He holds a Ph.D. in communication theory and research from the Annenberg School for Communication at the University of Southern California.

tp://www.nap.edu/catalog/12639.html

Fred A. Mettler, Jr., is a professor emeritus in the Department of Radiology at the University of New Mexico School of Medicine. He was chair of the department for 18 years, from 1984 to 2003. He is currently chief of Radiology and Nuclear Medicine at the New Mexico Federal Regional Medical Center. He is an academician of the Russian Academy of Medical Sciences and a Fellow of both the American College of Radiology and the American College of Nuclear Physicians. Dr. Mettler has authored more than 310 scientific publications, including 18 books on medical management of radiation accidents, medical effects of ionizing radiation, and radiology and nuclear medicine. He holds four patents. He was a scientific vice president of the National Council on Radiation Protection and Measurements, and remains a member. He has chaired two committees for the IOM and NRC. He is currently the U.S. Representative to the United Nations for Radiation Effects and is an emeritus commissioner of the International Commission on Radiation Protection. He was the health effects team leader of the International Chernobyl Project. He graduated with a B.A. in mathematics from Columbia University and an M.D. from Thomas Jefferson University. He received an M.P.H. from Harvard University.

Franklin E. Mirer is professor of Environmental and Occupational Health at Hunter College of the City University of New York. Previously, he served for decades as director of the Health and Safety Department for the United Automobile, Aerospace and Agricultural Implement Workers of America. His primary scientific interest is exposure and risk assessment in the occupational environment, and the interaction of science and policy in setting and enforcing health regulations. Dr. Mirer has served on National Academies committees on Institutional Means for Risk Assessment, Risk Assessment Methodology, and the Review of the Health Effects Institute. He has testified before House and Senate committees on occupational safety and health matters. He was inducted into the National Safety Council's Health and Safety Hall of Fame and is a Fellow of the *Collegium Ramazzini* and the American Industrial Hygiene Association. He holds appointments as an adjunct professor at the Michigan School of Public Health, adjunct associate professor at the Mt. Sinai School of Medicine, and visiting lecturer at the Harvard School of Public Health. He holds a Ph.D. in physical organic chemistry from Harvard University and is a toxicologist and certified industrial hygienist.

tp://www.nap.edu/catalog/12639.html

Jacqueline Nowell is director of the Occupational Safety and Health Office at the United Food and Commercial Workers International Union. Ms. Nowell and her staff develop and monitor ergonomic programs in the red meat, poultry, and retail industries. They develop educational materials and conduct training programs for local union stewards and leadership on a variety of safety and health issues in the union's represented industries. She is a member of the American Public Health Association and American Industrial Hygiene Association, and she serves on the NIOSH/National Occupational Research Agenda Traumatic Injuries and Special Populations at Risk Teams. She is currently a board member on the District of Columbia Occupational Safety and Health Board, which establishes policies related to occupational safety and health issues in the District of Columbia. She has worked for the New York Committee for Occupational Safety and Health and was an assistant professor at Hunter College's School of Health Sciences and Environmental and Occupational Health Science Program. Ms. Nowell received her M.P.H. from UCLA and is a certified industrial hygienist.

Raja V. Ramani is emeritus George H., Jr., and Anne B. Deike Chair of Mining Engineering and professor emeritus of mining and geoenvironmental engineering at The Pennsylvania State University. His research activities include mine health, safety, productivity, environment, and management; flow mechanisms of air, gas, and dust in mining environs; and innovative mining methods. Dr. Ramani has been a consultant to the United Nations, The World Bank, and the National Safety Council. He was the 1995 president of the Society for Mining, Metallurgy, and Exploration. He served on the U.S. Department of Health and Human Services' Mine Health Research Advisory Committee. He was chair of the National Academy of Sciences (NAS) Committee on Post Disaster Survival and Rescue and a member of the Health Research Panel of the NAS Committee on the Research Programs of the U.S. Bureau of Mines. He was a member of the Department of the Interior's Advisory Board to the Director of the U.S. Bureau of Mines and a member of the Secretary of Labor's Advisory Committee on the Elimination of Coal Worker's Pneumoconiosis. More recently, he was a member of the NAS committees on Technologies for the Mining Industries and on Coal Waste Impoundments. He also chaired the NRC–IOM Committee to Review the NIOSH Mining Safety and Health Research Program. Dr. Ramani is a member of the National Academy of Engineering. Dr. Ramani holds an

tp://www.nap.edu/catalog/12639.html

M.S. and a Ph.D. in mining engineering from Penn State, where he has been on the faculty since 1970.

Jorma Rantanen is the director emeritus of the Finnish Institute of Occupational Health. Dr. Rantanen has served as president of the International Commission on Occupational Health. He has led efforts to anticipate emerging workplace hazards, built an international network of occupational safety and health professionals, and improved working conditions in developing nations. He has been a pioneer in the development and recognition of comprehensive occupational health, including development of healthy and safe work environments, promotion and maintenance of work ability, and introduction of healthy work practices and lifestyles. He was awarded the NIOSH Lifetime Achievement Award in Occupational Safety and Health. He is the author of more than 430 research reports and book chapters covering medical biochemistry, radiation biology, toxicology, and risk assessment. Dr. Rantanen holds a Ph.D. in radiation biology and medical biochemistry from the University of Turku.

Richard L. Tucker is vice president of Tucker and Tucker Consultants, Inc., a private consulting organization related to management of large projects. Dr. Tucker's career has included three overlapping, but complementary, aspects. He has been a professor and researcher at two universities and, in retirement, holds the Joe C. Walter, Jr., Chair Emeritus at the University of Texas–Austin. He has been an academic administrator and director of a major research center involving several hundred construction industry executives. He has also had considerable direct industry involvement as an employee, an executive, and a consultant with several of the world's largest companies and projects. He has been personally involved in developing improved construction productivity methods, industry benchmarking, and metrics. Dr. Tucker has published more than 100 reports and journal articles. He is a member of the National Academy of Engineering and the National Academy of Construction. Dr. Tucker has a Ph.D. in civil engineering from the University of Texas.

James J. Zuiches is vice chancellor for extension, engagement, and economic development at North Carolina State University. He served previously as professor and extension specialist in the Department of Community and Rural Sociology at Washington State University (WSU).

tp://www.nap.edu/catalog/12639.html

Dr. Zuiches has also served as dean of the WSU College of Agricultural, Human, and Natural Resource Sciences; director of the Agricultural Research Center; director of Cooperative Extension; and vice chair and member of the Washington State Board of Natural Resources. His research specializations include demography, community and rural sociology, and research administration. He has also written on research priorities, indirect costs, distance education, and the returns on investment in research. Dr. Zuiches has received the Sustained Superior Performance Award from the NSF and is a fellow of the American Association for the Advancement of Science. He served on the NRC Committee on the Future of Colleges of Agriculture in the Land-Grant University System, and on the NRC Board on Agriculture and Natural Resources. He has a Ph.D. in sociology from the University of Wisconsin–Madison.

PREPUBLICATION COPY—UNCORRECTED PROOFS